

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2004, 07:50:32 ; Search time 8804 Seconds  
(without alignments)  
11124.144 Million cell updates/sec

Title: US-10-085-198-47  
Perfect score: 2071  
Sequence: 1 cccccccgccccgacccg.....ctatcacgtccctgttcc 2071

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_on.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2071	100.0	2071	6	AX921054 Sequence
2	1634.8	78.9	3457	9	BC041772 Homo sapi
3	1525.8	73.7	3179	6	AX179748 Sequence
4	1349.8	65.2	3273	10	BC011467 Mus muscu
5	1161.4	56.1	1353	6	CQ725703 Sequence
6	692.2	33.4	2342	6	BD242859 Secreted
7	563.8	27.2	1587	6	AX481748 Sequence
8	563.8	27.2	1836	9	BC006156 Homo sapi
9	563.8	27.2	2017	9	BC038961 Homo sapi
10	563.8	27.2	2049	9	BC065235 Homo sapi
11	563.8	27.2	2050	9	BC008325 Homo sapi
12	563.8	27.2	2165	9	AF212371 Homo sapi
13	563.8	27.2	2230	6	AX481746 Sequence
14	542.8	26.2	2201	6	BD191275 186 human
15	542.8	26.2	2201	6	AX924785 Sequence
16	532	25.7	2171	10	BC026854 Mus muscu
17	530.4	25.6	1731	10	BC002297 Mus muscu
18	524	25.3	2211	10	AF212372 Mus muscu
19	512.4	24.7	1661	6	BD191276 186 human

20	512.4	24.7	1661	6	AX924786	Sequence
21	453	21.9	1460	6	BD191122	186 human
22	453	21.9	1460	6	AX924632	Sequence
23	452.4	21.8	2614	5	BC048024	Danio rer
24	449.2	21.7	2643	5	AF465772	Danio rer
25	437.8	21.1	1915	9	AK172832	Homo sapi
26	437	21.1	112130	2	AC135621	Homo sapi
27	437	21.1	165762	2	AC124263	Homo sapi
28	437	21.1	178714	2	AC129504	Homo sapi
29	437	21.1	185378	9	AC118754	Homo sapi
30	437	21.1	190466	2	AC022285	Homo sapi
31	437	21.1	199463	2	AC129497	Homo sapi
32	434.6	21.0	1539	6	AX328008	Sequence
33	434.6	21.0	1897	6	AX709979	Sequence
34	434.6	21.0	2121	6	AX328006	Sequence
35	428.8	20.7	154622	2	AC126336	Homo sapi
36	399	19.3	56981	9	AC126391	Homo sapi
37	399	19.3	165762	2	AC124263	Homo sapi
38	399	19.3	168070	2	AC126328	Homo sapi
39	399	19.3	178714	2	AC129504	Homo sapi
40	399	19.3	180317	2	AC025590	Homo sapi
41	399	19.3	185378	9	AC118754	Homo sapi
42	399	19.3	190466	2	AC022285	Homo sapi
43	399	19.3	199463	2	AC129497	Homo sapi
44	388	18.7	35758	2	AC127539	Homo sapi
45	368	17.8	1431	6	AX191614	Sequence

ALIGNMENTS

RESULT 1  
AX921054  
LOCUS AX921054 2071 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 47 from Patent WO02068652.  
ACCESSION AX921054  
VERSION AX921054.1 GI:40214759  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Nov-x proteins and nucleic acids encoding same  
TITLE Patent: WO 02068652-A 47 06-SEP-2002;  
JOURNAL Location/Qualifiers  
FEATURES  
source  
1..2071  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2071; DB 6; Length 2071;  
Best Local Similarity 100.0%; Pred. No. 1.7e-290;  
Matches 2071; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CCCCCCGCCCGATCCCGGCGGCATGATGCTGGAATCGCGCTCGGCGGCGCG 60  
Db 1 CCCCCCGCCCGATCCCGGCGGCATGATGCTGGAATCGCGCTCGGCGGCGCG 60  
Qy 61 GCGCGCGCGAGGAGGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAG 120  
Db 61 GCGCGCGCGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAG 120  
Qy 121 CGAGGGGCTGGCGGTAGCGGTTCGCGGGCGCGGGGCGCGGGCGCGGTGGAGTCTCG 180  
Db 121 CGAGGGGCTGGCGGTAGCGGTTCGCGGGCGCGGGGCGCGGGCGCGGTGGAGTCTCG 180  
Qy 181 GCGCGGGCGATAGGTGCGAGCGTGTGCGGCGAGGTAAAGCGGGGCGGCGGACCC 240  
Db 181 GCGCGGGCGATAGGTGCGAGCGTGTGCGGCGAGGTAAAGCGGGGCGGCGGACCC 240









```
Qy 1586 TCGCCACTGCGCTCTTCTTCTGTCAGCGACCGCGCGAGGCTGAGCAGCACCTG 1638
Db 1501 TCGCCACTGCGCTCTTCTTCTGTCAGCGACCGCGCGAGGCTGAGCAGCAGTG 1553

RESULT 4
BC011467
LOCUS BC011467 3273 bp mRNA linear ROD 16-SEP-2003
DEFINITION Mus musculus cDNA sequence BC011467, mRNA (cDNA clone
IMAGE:3154539), partial cds.
ACCESSION BC011467
VERSION BC011467.1 GI:15079261
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3273)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, K.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2238257
MEDLINE 12477932
PUBMED 2 (bases 1 to 3273)
REFERENCE Strausberg, R.
Direct Submission
Submitted (30-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapsb@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 23 Row: b Column: 10.
Location/Qualifiers
1..3273
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"

FEATURES
source
```

## gene

```
/clone="IMAGE:3154539"
/tissue_type="Mammary tumor, Brca1-/fl; MMTV-Cre model. 10
months old, gross tissue."
/clone_lib="NCI CGAP_Mam3"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6"
<1..3273
/genes="BC011467"
/notes="synonym: MGC37865"
/db_xref="LocusID:216892"
/db_xref="MGI:2384936"
<1..1775
/genes="BC011467"
/codon_start=3
/product="BC011467 protein"
/protein_id="AAH11467.1"
/db_xref="GI:15079262"
/db_xref="LocusID:216892"
/db_xref="MGI:2384936"
/translations="DGFGRDGAARAAHRAHLSGPGGLASPEPPAPPAAPARAMMC
LECAARAAAGABEABEADAEARRRRRGAQPGAGGACCGARGVGAGVVSADAEVQTLS
GVSRRVPSGLPSTPTGCAAAAKPSAPQPKASLGRGAAAIISLGNLVNLDL
YTAVLLDIQHPGVKDRGAGLLQSVFICSPWVAIPFGYLDGRFRKRWILSCGIFP
WSAVTSSSFIPQIYFWLLVLSRGLVGIGEASYSYTIPTIIGDITKNTKRLMSVYF
FAIPLGSGLYITGSSVQAAGDMHWRVSPVLMITGTLILVLPATKRGHADQLG
GQKARTSLWRDKALIRNRYVFSLSATSAVSATGALRWIPLYLHRAQVVGKTAI
TCNSPPGAKDSLIFGATCTGTGLVVTGATRCWLRTQADPLVCAVQVQKTAI
FICLIFAAKTSIVGAVICIFVGETLLFSNWTADILMYVVIPTREATAVALQSPTS
HLLDGAGSPYLIGFISDLIRQSKDSPLWEFLSLGYALMLCPFFVVLVGMFFLATALF
FLSDRAAEQQVNLVMPVASVKV"
misc_feature 549..914
/genes="BC011467"
/notes="sugar tr; Region: Sugar (and other) transporter"
/db_xref="CDB:pfam00083"
```

## ORIGIN

```
Query Match 65.2%; Score 1349.8; DB 10; Length 3273;
Best Local Similarity 85.6%; Pred. No. 4.3e-186;
Matches 1501; Conservative 0; Mismatches 252; Indels 0; Gaps 0;

Qy 1 CCCCCCGCCCGCCGATCCCGCGCGGCGGATGATGCTGGAATCGGCTCGCGCGCGCG 60
Db 99 CCCCCCGCCCGCCCGCGCGCGCGCGCGGCGGATGATGCTGGAATCGGCTCGCGCGCGCG 158

Qy 61 GCGCGCGCGCGAGGAGGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 159 GCGCGCGCGCGAGGAGGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 218

Qy 121 CGAGGGGCTCGCGGTAGCGGTTCGTCGCGGGCGCGGGCGCGCGCGCGCGCGCGCTCG 180
Db 219 CCAGGGGCTGAGGTAGCGGTTCGTCGCGGGCGCGGGGCTGTTGGGGGCGCTGGAGTCTG 278

Qy 181 GCGCGCGCGGATGAGGTGACAGCGTGTGCGGCGAGCGGTAAAGCGGGGCGCGACCGGACCC 240
Db 279 TCGCGCGGACGAGAGGTGACAGCGTGTGTCGCGGAGCGGTGAGGGGGGCTCCGCTCGGAGTC 338

Qy 241 CCGCGCACCCCGCGCACCCCGCGCTCGGAGCTACTGCAAGGGGCGCGCGCGCTCACGAG 300
Db 339 CCAGCATCCCCAGCACCCCGCTGCTGCGCAGCAGCTGCCAAGGGGCGCGCGCGCGCGAG 398

Qy 301 CCCAAACCGGCCAGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCATCTCTCAGCTTGGGC 360
Db 399 CCCAAACCGGCCAGCTTAGGCCGAGGGGGGGGCGAGCTGCTGCCATCTCTGAGTTAGGC 458

Qy 361 AACGTGCTCAACTACTTGGACAGGTACCGTGGCAGGCGCTCTTCTTGGACATCCACGAG 420
Db 459 AACGTGCTCAACTACTTGGACAGGTACCTGTGGCAGGTGTTCTTCTTGGACATCCACGAG 518

Qy 421 CACTTTGGGTCAAGGACCGGCGCGCGCTCTGTCAGTCAGTGTTCATCTGTAGTTC 480
Db 519 CACTTTGGGTCAAGGACCGGCGCGCGCTTGTGTCAGTCAGTGTTCATCTGTAGTTC 578

Qy 481 ATGGTGGCTGCCCCCATCTTTCGGCTACCTGGGCGACCGCTTCAACAGGAAGGTGATTCTC 540
```

Db 579 ATGGTGGCTGCCCCCATCTTTGGCTACTCTGGCGACCGCTTCATAGAAAGGTGATCTC 638  
Qy 541 AGCTGGCGCATTTCTTTGGTGGCGGTACCTTTCTCCAGCTCTTTCATATCCCGAGCAG 600  
Db 639 AGCTGGCGCATCTTCTTCTGGTCTGTGTCACTTCTCGAGTCTTTTCATCCCTCAGCAG 698  
Qy 601 TACTTCTGGCTGCTGTCTGTCTCGGGGGCTGGTGGGCATCGGGAGGCGAGCTACTCC 660  
Db 699 TACTTCTGGCTACTAGTCTGTCTCGAGGGCTAGTAGGTATTTGGTGAAGCCAGCTACTCC 758  
Qy 661 ACCATCGCCCCCACTATATCTGCGACCTTTTACCAGAACACACGCTACGCTCATGCTG 720  
Db 759 ACCATAGCACCCACCATCATTTGGCGACTTGTTTACCAAGAACACACGCACTGATGCTA 818  
Qy 721 TCCGCTCTTACTTGGCCATCCCACTGGCGAGTGGCTGGGCTACATTAATGCGCTCCAGC 780  
Db 819 TCTGTCTTCTATTTTGGCCATCCCCCTGGCGAGTGGCTGTATATACAGGTTCCAGC 878  
Qy 781 GTGAAGCAGGACCGCGAGACTGGGCATCGGGCATTTGCGGGTGTCCGCTGTCTGGGCATG 840  
Db 879 GTGAAGCAGGACGCTGGAGACTGGCATTTGGGGCCCTGCGGGTTTCCCGCTCTGGGCATG 938  
Qy 841 ATCAGAGGAACACTCATCTCTATTTGTGTCAGCCACTTAAAGGGGTCAATCCGACACAG 900  
Db 939 ATCAGAGGAACACTCATCTCTATCTTGTGTCAGCCACTTAAAGGGGTCAATCCGACACAG 998  
Qy 901 CTCGGGACAGCTCAAGGCCCGGACCTCATGGCTCCGAGATATGAAGGCCCTGTATCGA 960  
Db 999 CTTGGGGGCGAGCTCAAGCAGCGACCTCTGGCTTCGAGACATGAAGGCCCTGTATCGA 1058  
Qy 961 AACCGCAGTACGTCTTCTCTCTGCGCCAGCTGGCTGTCTCTTCCCGACGGGGGC 1020  
Db 1059 AACCGCAGTACGTCTTCTCTCTGCGCCAGCTGGCTGTCTCTTCCCGACGGGGGC 1118  
Qy 1021 CTGGGATGTGATCCGCTCTACCTGTGACCGCGCCAAAGTTGTGACAGAGCAGCAGAG 1080  
Db 1119 CTGGGATGTGATCCGCTCTATCTTCACTGCTCAAGTTGTGACAGAGCAGCAGAG 1178  
Qy 1081 AGCTGCAACAGCCGCGCTGTGGGGCCAAAGCAGCCTCATCTTTGGGGCCATCACTGC 1140  
Db 1179 ACTTGAACAGCCGCGCTGTGGAGCAAAGCAGCCTCATCTTTGGAGCCATTACTGC 1238  
Qy 1141 TTTACGGGATTTCTGGGCTGTGTACGGGGGACAGGACACCGGCTGGTGGCGCTGAAG 1200  
Db 1239 TTTACTGGCTTCTGGGCGTAGTACAGGGCGGGAGCCACTCGTTGGTGGCGCTGGCG 1298  
Qy 1201 ACCCAGGGCGGACCACTGGTGTGCGGTGGGCATCTGGGCTCTGCCATCTTTCATC 1260  
Db 1299 ACTCAACGTGCTGACCCATTTGGTGTGCTGTGGGCATCTGGGATCCGCGCATCTTCATC 1358  
Qy 1261 TGCTGATCTTCTGGCTGCGCAAGACAGCAGCATCGTAGGAGCCATATCTGTATCTTCGTC 1320  
Db 1359 TGCTCATCTTTGTGCTGCCAAGACAGCATTTGTAGGCGCTTATATCTGCATTTTGT 1418  
Qy 1321 GGGGAGACGCTGTGTTTCTAACTGGGCGCATCACTGCGAGACATCTCATGTATGATGCTG 1380  
Db 1419 GGAGAGACATGCTGTTTCTTAACTGGGCGCATCACTGCGAGACATCTCATGTATGCTG 1478  
Qy 1381 ATCCCAACGGGGCGGCGCACTGCGGTGGCTTGCAGAGCTTCACTCCCGCTGCTGGG 1440  
Db 1479 ATCCCACTGCGGAGCGCACTGCTGTGGGCGCTGCGAGCTTCACTCCCATCTGCTGGG 1538  
Qy 1441 GAGCGCGGAGCGCCCTACCTCATTTGGCTTTTATCTCAGACCTGATCCGCCAGAGCACTAAG 1500  
Db 1539 GAGCTGGAAGCCCTTACCTCATTTGCTTATCTCGAACCTTATCCGCCAGAGCACTAAG 1598  
Qy 1501 GACTCCCGCTCTGGAGTTCTGAGCCTGGGCTACGCGCTCATGTCTGCGCTTTCGTC 1560  
Db 1599 GACTCCCGCTCTGGAGTTCTGAGCCTGGGCTATGCCCTCATGTGTGCCCTTTGTC 1658  
Qy 1561 GTGGTCTGGGGGATGTTCTTCTCGCACCTGCGCTCTTCTGTCAGCGACCGCGC 1620

Db 1659 GTGGTCTGGGTGGCATGTTCTTCTTCCACTGCTCTCTTCTTCTCAGCGACCGTGC 1718  
Qy 1621 AGGCTGAGCAGCACCTGGGGAGAGACGGGGGGTTCAGGGTGGTGCATCAGCGGGG 1680  
Db 1719 AAGCTGAGCAGCAGGTGAACAGCTGGTGTATGCTCTCGATCCGTGAAAGTCTGAGT 1778  
Qy 1681 CGGGCCCGGCACTGCTTCTGGCATCTGTCTGGGGGGCCAGCTGACCGAGGTGCTG 1740  
Db 1779 GGTGCGGTGGACAGCAATGAATGCACTGCACTTGTAGTTGGAGAGGTCTTCCAGCATCTCTG 1838  
Qy 1741 GGCAGGGACCTCG 1753  
Db 1839 GACCGCAGGACG 1851

RESULT 5  
LOCUS CQ725703 1353 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 11637 from Patent WO02068579.  
ACCESSION CQ725703  
VERSION CQ725703.1 GI:42287032  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.  
AUTHORS Kite, such as nucleic acid arrays, comprising a majority of  
TITLE humanexons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 11637 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES Location/Qualifiers  
source 1..1353  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 56.1%; Score 1161.4; DB 6; Length 1353;  
Best Local Similarity 99.5%; Pred. No. 9.5e-159;  
Matches 1165; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 468 CATCTGTAGCTTCATGTGGCTGCCCCCATCTTCGGCTACCTGGGCGCGCTTCAACAG 527  
Db 144 CGTTGGAAGCTTCATGGTGGCTGCCCCCATCTTCGGCTACCTGGGCGCGCTTCAACAG 203  
Qy 528 GAAGGTGATTTCTCAGCTGGGCATTTTCTTCTGGTGGGCGCTCACCTTCTCCAGCTCTT 587  
Db 204 GAAGGTGATTTCTCAGCTGGGCATTTTCTTCTGGTGGGCGCTCACCTTCTCCAGCTCTT 263  
Qy 588 CATTTCCCGCAGCTACTTCTGGCTGCTGCTCTGCTCCCGGGGGCTGGTGGGCATCGGGGA 647  
Db 264 CATTTCCCGCAGCAGTACTTCTGGCTGCTGCTCTGCTCCCGGGGGCTGGTGGGCATCGGGGA 323  
Qy 648 GCGCAGCTACTCCACATCGCCCCCATCTATCATTTGGCGACCTTTCACCAAGAACACGG 707  
Db 324 GCGCAGCTACTCCACATCGCCCCCATCTATCATTTGGCGACCTTTCACCAAGAACACGG 383  
Qy 708 TACGCTCATGCTGCTCGCTTCTTACTTCGCCATCCACATCGGCGAGTGGGCTGGGCTACAT 767  
Db 384 TACGCTCATGCTGCTCGCTTCTTACTTCGCCATCCACATCGGCGAGTGGGCTGGGCTACAT 443  
Qy 768 TACTGGCTCCAGCTGAAGCAGCAGCGAGACTGGCACTGGGCATTCGGGCTGTCCCC 827  
Db 444 TACTGGCTCCAGCTGAAGCAGCAGCGAGACTGGCACTGGGCTGTCCCC 503  
Qy 828 TGTCTGGGCATGATCAGAGAACACTCATCTCTTCTGTCGCCGACCTTAAAAAGGG 887  
Db 504 TGTCTGGGCATGATCAGAGAACACTCATCTCTTCTGTCGCCGACCTTAAAAAGGG 563  
Qy 888 TCATGCGCAGCTCGGGGACAGCTCAAGGCCCGGACCTCATGCTCCGAGATATGAA 947

564	Db	TCATGCCGACAGCTCGGGACACAGCTCAAGGCCGAGCTCTCATGGCTCGAGATATGAA	623
948	Qy	GGCCCTGATTGGAACCGCACTAGCTCTTCTCTCCTGCGCCACGTCGGCTGCTCTCCTT	1007
624	Db	GGCCCTGATTGGAACCGCACTAGCTCTTCTCTCCTGCGCCACGTCGGCTGCTCTCCTT	683
1008	Qy	CGCCACGGGGCCCTGGGCATGTGGATCCCGCTCTACTGTCACCGCGCCCAAGTTGTGCA	1067
684	Db	CGCCACGGGGCCCTGGGCATGTGGATCCCGCTCTACTGTCACCGCGCCCAAGTTGTGCA	743
1068	Qy	GAAGACAGCAGACGTCGCAACAGCCGCCCTGTGGGGCCCAAGGACAGCCTCATCTTTGG	1127
744	Db	GAAGACAGCAGACGTCGCAACAGCCGCCCTGTGGGGCCCAAGGACAGCCTCATCTTTGG	803
1128	Qy	GGCCATCACCTGCTTTACGGGATTTCTGGGGTGGTCA CGGGGGCAGGACGACGCGCTG	1187
804	Db	GGCCATCACCTGCTTTACGGGATTTCTGGGGTGGTCA CGGGGGCAGGACGACGCGCTG	863
1188	Qy	GTCCGGCTCAAGACCCAGCGGGCCGACCACTGGTGTGTGCCGTGGGCATGCTGGGCTC	1247
864	Db	GTCCGGCTCAAGACCCAGCGGGCCGACCACTGGTGTGTGCCGTGGGCATGCTGGGCTC	923
1248	Qy	TGCCATCTTCATCTGCCCTGATCTTTCTGGGTGTCGCAAGACAGCATCGTAGGAGCTATAT	1307
924	Db	TGCCATCTTCATCTGCCCTGATCTTTCTGGGTGTCGCAAGACAGCATCGTAGGAGCTATAT	983
1308	Qy	CTGTATCTTCGTGGGGAGACGTCGCTGTTTTCTTAACTGGGCCATCACTGGACATCCT	1367
984	Db	CTGTATCTTCGTGGGGAGACGTCGCTGTTTTCTTAACTGGGCCATCACTGGACATCCT	1043
1368	Qy	CATGTACGTGTCATCCCCACGGGGCGGCACCTGCCGTGGCCTTGACAGAGTTCACTC	1427
1044	Db	CATGTACGTGTCATCCCCACGGGGCGGCACCTGCCGTGGCCTTGACAGAGTTCACTC	1103
1428	Qy	CCACCTGCTGGGGACGCCGGAGCCCTCACTCATTTGGCTTTATCTCAGACCTGATCCG	1487
1104	Db	CCACCTGCTGGGGACGCCGGAGCCCTCACTCATTTGGCTTTATCTCAGACCTGATCCG	1163
1488	Qy	CCAGAGCACTAAGACTCCCGCTCTGGGAGTTCTTGAGCCTGGGCTACCGGCTCATGCT	1547
1164	Db	CCAGAGCACTAAGAGACTCCCGCTCTGGGAGTTCTTGAGCCTGGGCTACCGGCTCATGCT	1223
1548	Qy	CTGCGCTTTCTGCTGGTCTCTGGGGGCACTGTTCTTCTCTCGCCACTGCGCTCTTCTCGT	1607
1224	Db	CTGCGCTTTCTGCTGGTCTCTGGGGGCACTGTTCTTCTCTCGCCACTGCGCTCTTCTCGT	1283
1608	Qy	CAGCGACCGCGCCAGGGCTCAGACGACCTTG	1638
1284	Db	CAGCGACCGCGCCAGGGCTCAGCAGCAGGTG	1314

RESULT 6	BD242859	LOCUS	BD242859	DEFINITION	Secreted proteins and polynucleotides encoding them.	2342 bp	DNA	linear	PAT 17-JUL-2003
DESCRIPTION	BD242859	ACCESSION	BD242859	VERSION	1	GI:33052629			
KEYWORDS	JP 2002536973-A/10.								
SOURCE	Homo sapiens (human)								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
AUTHORS	1 (bases 1 to 2342)								
TITLE	Valenzuela,D., Yuan,Q., Hoffman,H., Hall,J. and Rapiejko,P.								
JOURNAL	Secreted proteins and polynucleotides encoding them								
	Patent: JP 2002536973-A 10 05-NOV-2002;								
	ALPHAGEN INC								
COMMENT	OS Homo sapiens (human)								
	PN JP 2002536973-A/10								
	PD 05-NOV-2002								
	PF 18-FEB-2000 JP 2000599860								
	PR 19-FEB-1999 US 60/120680, 23-APR-1999 US 09/298733 PR								

17-AUG-1999	US	60/149639,23-SEP-1999	US	60/155686	PR
01-OCT-1999	US	60/157247,29-NOV-1999	US	60/167823	PR
29-NOV-1999	US	60/167822,15-FEB-2000	US	60/162711	PI
VALENZUELA,OLIVE	YUAN,HEIDI	HOFFMAN,JEFF	HALL,PIETER	PI	RAPJEJKO
PC	C12N15/09,A61K38/00,A61P3/10,A61P5/14,A61P11/00,A61P11/06,PC				
A61P19/02,					
PC	A61P21/04,A61P25/14,A61P27/02,A61P29/00,A61P31/04,A61P31/10,				
PC	A61P31/12,				
PC	A61P31/18,A61P31/20,A61P31/22,A61P37/00,A61P37/06,C07K14/435,				
PC	C12N5/10,				
PC	C12P19/34/(C12P19/34,C12P1:91),C12N15/00,C12N5/00,A61K37/02				
CC	Secreted proteins and polynucleotides encoding them	PH	Key		
Location/Qualifiers					
FT	source	1..2342			
FT	Location/Qualifiers	/organism='Homo sapiens (human)'			
	1..2342				
/organism='Homo sapiens'					
/mol_type='genomic DNA'					
/db_xref='taxon:9606'					
ORIGIN					
Query Match	33.4%;	Score 692.2;	DB 6;	Length 2342;	
Best Local Similarity	99.8%;	Pred. No. 7.6e-91;			
Matches 694;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;	
QY	942	TATGAAGGCCCTGATTTCGAAACCGCAGCTACGTCTTCTCCTCCCTGGCCACAGCTCGGCTGT	1001		
DB	1	TATGAAGGCCCTGATTTCGAAACCGCAGCTACGTCTTCTCCTCCCTGGCCACAGCTCGGCTGT	60		
QY	1002	CTCCTTTGCCACCGGGGGCCCTGGGCATGTGGATCCCGCTCTACCTGTCACCGGCCCAAGT	1061		
DB	61	CTCCTTTGCCACCGGGGGCCCTGGGCATGTGGATCCCGCTCTACCTGTCACCGGCCCAAGT	120		
QY	1062	TGTGCAGACAGCAGACAGCTGCACAGCCGCCCTGTGGGGCCCAAGCAGCCTCAT	1121		
DB	121	TGTGCAGACAGCAGACAGCTGCACAGCCGCCCTGTGGGGCCCAAGCAGCCTCAT	180		
QY	1122	CTTTGGGGCCCATCACCTGCTTTACGGGATTTCTGGGCGTGTTCACGGGGCAGGAGCCAC	1181		
DB	181	CTTTGGGGCCCATCACCTGCTTTACGGGATTTCTGGGCGTGTTCACGGGGCAGGAGCCAC	240		
QY	1182	GGCTGGTCCGCTGAAACCCAGCGGCCCGACCCATCTGTGTGTGCGGTGGGCATGCT	1241		
DB	241	GGCTGGTCCGCTGAAACCCAGCGGCCCGACCCATCTGTGTGTGCGGTGGGCATGCT	300		
QY	1242	GGGCTCTGCATCTTCATCTGCTGATCTTCTGTGGTGCACAGAGCAGCATCTGAGGAC	1301		
DB	301	GGGCTCTGCATCTTCATCTGCTGATCTTCTGTGGTGCACAGAGCAGCATCTGAGGAC	360		
QY	1302	CTATATCTGTATCTTCGTGGGGAGACGCTGCTGTTTTCTAACTCGGCCATCACATGCAGA	1361		
DB	361	CTATATCTGTATCTTCGTGGGGAGACGCTGCTGTTTTCTAACTCGGCCATCACATGCAGA	420		
QY	1362	CATCTCATGTAGTGGTTCATCCCAACCGCGGCCCATCTGCCGTGGCCCTTCAGAGCTT	1421		
DB	421	CATCTCATGTAGTGGTTCATCCCAACCGCGGCCCATCTGCCGTGGCCCTTCAGAGCTT	480		
QY	1422	CACCTCCACCTGTCTGGGGAGCGCGGAGACCCCTTACTCATTTGGCTTTATCTCAGACT	1481		
DB	481	CACCTCCACCTGTCTGGGGAGCGCGGAGACCCCTTACTCATTTGGCTTTATCTCAGACT	540		
QY	1482	GATCGCCAGACACTTAGGACTCCCGCTCTGGGAGTTCCTGAGCCTGGGCTAGCGCT	1541		
DB	541	GATCGCCAGACACTTAGGACTCCCGCTCTGGGAGTTCCTGAGCCTGGGCTAGCGCT	600		
QY	1542	CATGCTCTCCCTTTCTGTCGTGCTCTGGCGGCATGTTCTTCTTCGACACTTCGCTCTT	1601		
DB	601	CATGCTCTCCCTTTCTGTCGTGCTCTGGCGGCATGTTCTTCTTCGACACTTCGCTCTT	660		
QY	1602	CTTGTGCAGCAGCCGCGCAGGGCTGAGCAGACCTG	1638		
DB	661	CTTGTGCAGCAGCCGCGCAGGGCTGAGCAGAGGTG	697		



SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Strausberg R.L., Reingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, K.H., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marudina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalley, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE	22388257
PUBMED	12477932
REFERENCE	2 (bases 1 to 1836)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (02-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgabbs-r@mail.nih.gov">cgabbs-r@mail.nih.gov</a> Tissue procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada <a href="mailto:info@bcgsc.bc.ca">info@bcgsc.bc.ca</a>
FEATURES	Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Len Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalley, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.
source	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAL Plate: 18 Row: h Column: 10. Location/Qualifiers 1..1836 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:3627317" /tissue_type="Brain, neuroblastoma" /clone_lib="NIH MGC_19" /lab_host="DH10B-R" /note="Vector: pOTB7" 1..1836 /gene="SPINL" /db_xref="LocusID:83985" 1..1575 /gene="SPINL" /codon_start=1 /product="SPINL protein"
gene	
CDS	







cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: nisc.mc@nhgri.nih.gov  
 Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakeley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
 Maduro, Q.L., Masiello, C., Maekari, B., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
 Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 140 Row: m Column: 14  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Genomescan gene  
 prediction, Similarity but not identity to protein.

## FEATURES

## source

1..2049  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:70841 IMAGE:6161778"  
 /tissue\_type="Skin, melanotic melanoma."  
 /clone\_lib="NIH MGC\_72"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"

## gene

1..2049  
 /gene="SPINL"  
 /db\_xref="LocusID:83985"  
 123..1709  
 /gene="SPINL"  
 /codon\_start=1  
 /product="SPINL protein"  
 /protein\_id="AAH65235.1"  
 /db\_xref="GI:40807118"  
 /db\_xref="LocusID:83985"  
 /translation="MAGSDTAPFLSQDDPDDGPVPGTGLPGSTNPKSERPEVPDQ  
 EGLRTGLSPGRSLIVAVLCVILNLYMDRFTVAGLPDIEQFENIGDSSGLIQT  
 VTFSSVMVLAPVGYLDGRYKYLKMGCIAPMSLVTLGSSPIRGEHFWLLLTGLV  
 GYGEASYSIATPLIADLVADQSRMLSIPIYPAIPVSGSLGYIAGSKVDNAGDWHV  
 ALRVTPLGVFVLLFLVVRPPRGVERHSDLPPLNPTSWADRLARLNPFSVL  
 SLGFTAVAVTGLALWAPAFLLRSRVILGETPPCLPGDSCSSDSLIPLGLITLGV  
 LGVGLGVEISRILSRHNPADPLVCATGILLGAPFLFLACARGSIYATYIFIFIGE  
 TLLSMWAIIVADILLYVLPTRRSIAEAFQIVLSHLGDAGSPYILGLISDLRLRNWP  
 PGFLSFRALQFLMLCAFVGLGGAFLGTALFIEADRRRAQLHVQGLLHEAGSTDD  
 RIVVPGRSTRVPVASVLI"

## CDS

393..776  
 /gene="SPINL"  
 /note="Sugar tr; Region: Sugar (and other) transporter"  
 /db\_xref="CDD:pfam00083"

## misc\_feature

393..776  
 /gene="SPINL"  
 /note="Sugar tr; Region: Sugar (and other) transporter"  
 /db\_xref="CDD:pfam00083"

## ORIGIN

Query Match 27.2%; Score 563.8; DB 9; Length 2049;  
 Best Local Similarity 64.3%; Pred. No. 3e-72;  
 Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;  
 QY 282 GGGCCCCGGCGCTCAGCAGCCCAACCGCCAGCTTGGCCGCGGGGGGCGAGCCGC 341  
 DB 248 GGACAGAGGGGCTGCAGGCATCACCGGCTGTCTCCGCGCGTTCGCGCTCATAGT 307  
 QY 342 CGGCATCTCAGCTTGGGCAACGCTGCTCAACTACCTGGACAGGTACACCGTGGCAGCGT 401  
 DB 308 GCGCGTGTGTGTACATCAATCTCTGAACTACATGGACCGCTTCAACGCTGGCGT 367  
 QY 402 CTTCTCGACATCCAGCAGCATTTGGGCTCAAGNACGAGGGCGCGCTCTCTCAGTC 461  
 DB 368 CTTTCCGACATCGAGCAGTTCTTCAACATCGGGGACAGTAGCTCTGGGCTCATCAGAC 427

QY 462 AGTGTTCATCTCTAGCTTCATGGTGGCTGCCCCCATCTTTCGGCTACCTGGGGGACGCGTT 521  
 DB 428 CGTGTTCATCTCCAGTTACATGGTGTGGCACCTGTGTGGTACCTACCTGGGTGACAGGTA 487  
 QY 522 CAACAGGAGGTGATTCTCAGCTGGGGCATTTTCTTGTGGTGGCGCGCTACCTTCTCCAG 581  
 DB 488 CAATCGGAAGTATCTCATGTGGGGGCATTTGCCCTTCTGGTCCCTGGTGACACTGGGGTC 547  
 QY 582 CTCTCTTCATTTCCCAAGCAGTACTTCTGGCTGTGTGCTCTCTCCCGGGGCTCTGGTGGCAT 641  
 DB 548 ATCTCTTCATCCCGGAGAGCATTTCTGGCTGTCTCTCTGACCCCGGGGCTCTGGTGGGT 607  
 QY 642 CCGGAGGCGCAGTACTCTCACCATCGCCCCCACTATCATTTGGCGACCTTCTTCCCAAGAA 701  
 DB 608 CGGGGAGGCCAGTTATTCCACCATCGCGCCCACTCTCATTTGCGGACCTCTTTGTGGCCGA 667  
 QY 702 CACGCGTAGCTCATGTCTGCTGCTTCTACTTTCGCCATCCCATCTGGGAGTGGCTGG 761  
 DB 668 CCAGCGGAGCGGATGCTCAGCATTTCTTACTTTGCCATTCGGTGGGCGAGTGGTCTGG 727  
 QY 762 CTACATTACTGGCTCCAGCGTGAAGCAGCGCAGCAGCTGGCAGCTGGGCAATTGGGGT 821  
 DB 728 CTACATTGACGGCTCCAAAGTGAAGGATATGCTGAGACTGGCACTGGGCTCTGAGGGT 787  
 QY 822 GTCCCTCTGTCTGGGCATGATCAGGAACTCATCTCTCATTTCTGGTCCCGAGCCACTAA 881  
 DB 788 GACACCGGGTCTAGGAGTGGTGGCGGTTCTGCTGTGTTCTCTGTGTGAGTGGCGGAGCCGC 847  
 QY 882 AAGGGGTCTATGCGCAGCAGCTCGGGGA--CCAGCTCAAGGCCCGGAGCTCATGTGCTCG 938  
 DB 848 AAGGGAGCGCTGGAGCGCCACTCAGATTGGCACCCCTGAACCCCACTCTGTGTGGGC 907  
 QY 939 AGATATGAAGCCCTGATTTCGAAACCGCAGTACGTCTTCTCTCTCTGGCCACGCTCGGC 998  
 DB 908 AGATCTAGGGCTCTGGCAGAGAAATCTTAGTTTGTCTCTGTCTTCTCTGGGCTTCACTGC 967  
 QY 999 TGTCTCTTTCGCCACGGGGGCTGGGCATGTGGATCCGGCTCTACTCTGACCGGCCCA 1058  
 DB 968 TGTGGCTTTTGTACGGGCTCCCTGGCTCTGTGGGCTCCGGCATTTCTGTGCTGTTCGCG 1027  
 QY 1059 AGTTGTGCA---GAAGACAGCAGAGACGTGCAACAGCCCGGCTGTGGGGCCAGGACAG 1115  
 DB 1028 CGTGTCTTGGGGAGACCCCACTGCTTCCCGGAGACTCTCTCTCTCTGACAG 1087  
 QY 1116 CTTATCTTTGGGGCATCACTGCTTTTACGGGATTTCTGGGCGTGTGTACCGGGGACAG 1175  
 DB 1088 TCTCATCTTTGGACTCATCACTGCTGACCGGAGTCTCTGGGTGTGGGCTCTGGGTGTGGA 1147  
 QY 1176 AGCCACGCGTGTGCGGCTGAAGACCCAGCGGGCGACCCACTGTGTGTGTGCGGTGG 1235  
 DB 1148 GATCAGCCCGGCTCCGCCACTCAACCCCGGGCTGATCCCTGTGTGTGTGCCACTGG 1207  
 QY 1236 CATGCTGGGCTCTGCGCATTTTCATCTGCTGATCTTCTGTGGCTGTCGAAGAGCAGCATCGT 1295  
 DB 1208 CTTCTGGGCTCTGCACCTTCTCTTCTCTCTCTGCTTGGCTGCGCCGCTGGTGTAGCATCGT 1267  
 QY 1296 AGGAGCTATATCTGTATCTTCTGCTGGGAGACCGTGTCTGTTTCTAACTGGGCCATCAC 1355  
 DB 1268 GGCCACTTATATTTTTCATCTTTTATGGAGAGACCTCTCTGTCATGAACCTGGGCCATCGT 1327  
 QY 1356 TGCAGACATCTCATGTACGTGGTTCATCCCGCGGGCGCCACTGCGGTGGCTTGTGCA 1415  
 DB 1328 GGCGACATTTCTGTGTACGTGTGTATCCCTTACCGAGCTTCCACCGCGGAGGCTTCCA 1387  
 QY 1416 GAGCTTCACTCCCACTGTGTGGGGAGCGCGGAGGCCCTTCACTCATTTGCTTTATCTC 1475  
 DB 1388 GATCGTGTGTCCACCTCTGCTGGGTGTGCTGGGAGGCCCTACCTCATTTGCGCTGATCTC 1447  
 QY 1476 AGACTGATCGCGCAGAGCAGTAAAGGACTCCCGCTCTGGGAGATTCTTGACCTGGGGTA 1535  
 DB 1448 TGACCGCTTGGCGGGAATGGGCCCTCTCTTCTTGTTCGAGTTTCGGGCTCTGCAAGT 1507



Db 868 GACACCGGGTCTAGGAGTGGTGGCGTTCTGCTGCTGTTCTCTGTTAGTGGGAGCCGCC 927  
QY 882 AAGGGGTCTAGCCAGACAGTCTGGGA----CAGCTCAAGGCCCGGACCTCTATGCTCCG 938  
Db 928 AAGGGAGCGGTGGAGCGCACTCAATTTGCGACCCCTGAACCCACCTCGTGGTGGC 987  
QY 939 AGATATGAAGGCCCTGATTGAAACCGCAGCTACGCTCTTCTCCCTCGGCGACGTCGGC 998  
Db 988 AGATCTGAGGGCTCTGGCAGAAATCTTAGTTTCGCTGCTGCTTCCCTCGGCTTCACTGC 1047  
QY 999 TGTCTCTTTCGCCACCGGGGCCCTGGGCATGTGATCCCGCTCTACCTGCACCCGCGCCA 1058  
Db 1048 TGTGGCTTTGTACCGGGCTCCCTGGCTGTGGGCTCGGCACTTCTGCTGCTTCCG 1107  
QY 1059 AGTTGTGCA--GAAGACAGACAGACGTGCAACAGCCCGCCCTGTGGGGCGAGCAG 1115  
Db 1108 CTGTGTCTTGGGAGACCCACCTGCTTCCCGGAGACTCTCTGCTTCTCTGTGACAG 1167  
QY 1116 CTCTATCTTTGGGCCATCAGCTGCTTTTACCGGATTTCTGGGCTGTGTGCGGGCGAG 1175  
Db 1168 TCTCATCTTTGAGCTCATCAGCTGCTGACCGGAGTCTTGGGTGTGGCTGGTGTGGA 1227  
QY 1176 AGCCAGCGCTGTGGCTGAGACCCAGCGGGCGGACCACTGTGTGTGTGCGGTGG 1235  
Db 1228 GATCAGCGCGGCTCGCCCACTCCAAACCCCGGGCTGATCCCTGCTGTGCACTGG 1287  
QY 1236 CATGTGGGCTCTGCATCTTTCATCTGCTGATCTTCTGCTGGCTGCCAAGACGACATCGT 1295  
Db 1288 CTCTCTGGGCTCTGCACCTTCTCTTCTGCTGCTTCTGCTGCGCCGCTGGTAGCATCGT 1347  
QY 1296 AGGAGCCTATATCTGTATCTTCTGCTGGGAGACGCTGCTGTTTTCTAACTGGGCCATCAC 1355  
Db 1348 GGCCACTTATATTTTATCTTCTATTGGAGAGACCTCTCTGTCATGAATGGGCCATCGT 1407  
QY 1356 TGCAGACATCTCATGTACGTGTCATCCCAACCGCGCGGCGCACTGCGTGGCTTGA 1415  
Db 1408 GGCCGACATCTCTGTGATGTTGATCCCTACCCGACGCTCCACCGCGGAGCCCTTCA 1467  
QY 1416 GAGCTTCACTTCCACTCTGCTGGGGAGCGCGGGAGCCCTTACTCTATTGGCTTTATCTC 1475  
Db 1468 GATCGTCTGTCCACCTGCTGGGTGATGCTGGAGCCCTTACTCTATTGGCTGATCTC 1527  
QY 1476 AGACCTCATCGCCAGACACTAAGNACTCCCGCTCTGGGAGTTCCTGAGGCTGGGCTA 1535  
Db 1528 TGACCGCTCGCGGAACTGGCCCCCTCTTCTTGTTCGAGTTCCGGGCTCTGAGTT 1587  
QY 1536 CGGCTCATGCTCTGCGCTTTCTGCTGTGTGCTTGGGCGGATGTTCTTCTCGCCACGTC 1595  
Db 1588 CTGCTCATGCTCTGCGCGTTTGTGGGCACTTGGCGGCGGAGCCCTTCTGGGCAACCGC 1647  
QY 1596 GCTCTTCTGTCAGCGCGCGCCAGGGGTGAGCAGCACTGGGGG 1642  
Db 1648 CATCTTCAITGAGCGCAGCGCGGCGGACAGCTGACGTGACGTGTCAGG 1694

RESULT 12  
AF212371  
LOCUS Homo sapiens spinster-like protein mRNA, linear PRI 21-MAY-2001  
DEFINITION AF212371  
ACCESSION AF212371  
VERSION AF212371.1 GI:12003979  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)

REFERENCE  
AUTHORS  
Shimoda, Y., Fujitani, K., Kurihara, J., Regan, J., Usui-Aoki, K.,  
Awano, W., Kaneda, M., Umeda, M. and Yamamoto, D.  
Mutations in the novel membrane protein spinster interfere with  
programmed cell death and cause neural degeneration in *Drosophila*  
melanogaster

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
SUBMISSION  
JOURNAL

Mol. Cell. Biol. 21 (11), 3775-3788 (2001)  
21238517  
11340170  
2 (bases 1 to 2165)  
Nakano, Y. and Yamamoto, D.  
Submitted (08-DEC-1999) Pioneer Research, Mitsuishi Kasei  
Institute of Life Sciences, 11, Minamioya, Machida, Tokyo  
194-8511, Japan

FEATURES  
source

1. 2165  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
325--1911  
/note="similar to Drosophila spinster protein; SPIN1"

CDS

/codon\_start=1  
/product="spinster-like protein"  
/protein\_id="AAG43830.1"  
/db\_xref="GI:12003980"  
/translation="MAGSDTAPFLSQADDDPDPVPGTGLPGSTGNPKSEEPVDDQ  
EGLORINGLSPPRSALIVAVLCYINILNMDRFTVAGVLEDFEQFNIGSSSLIQT  
VFISYVVLAPVFGILGDRNRYKMGCGIAFWSLVLTGSSFIPEHFHLLLRGLV  
GVGASSTIATLIADLFVADQSRMSLFYFAIPVSGGLGTAGSKVMDMADHWH  
ALRVTGGLVVAVLLFLVREPPRGAVRHSDLPPLNPTSMWADLRALRNPSFVLS  
SLGFTAVFTVSLALWAPALRSRVLTGTPCLPGDSCSSDSLSIFGLITLCTGV  
LGVGLVEISRRLRHSNPRADPLVCAATGLLGSAFPLFLSLACAGSIVATYPIFIGE  
TLLSMNVAIVADILLYVIVIPTRSTAEAFIVLSHLGLDAGSPVLTGLISDRLRNWP  
PSPLBFRALQFSIMLCFAVGLGGAFLGTAIFIEADRRRAQLHVQGLLHEAGSTDD  
RIVVPQGRSTRVPVASVLI"

ORIGIN

Query Match 27.2%; Score 563.8; DB 9; Length 2165;  
Best Local Similarity 64.3%; Pred. No. 3e-72;  
Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;

QY 282 GGGCCCCGGGCTCAGCAGCCCAACCGCGCAGCTTGGCGCGGGGGGGGAGCCGC 341  
Db 450 GGACGAGAGGGGTGCGCGGCATCACCGGCTGTCTCCCGGCGCTTCGCTCTCATAGT 509  
QY 342 CGCCATCTCAGCTTGGGCAACGCTCAACTCTGGACAGGTACACGTTGGCAGCGCT 401  
Db 510 GCGGTGCTGTGTACATCACTCTCTGAACATACGACCGCTTACCGTGGCTGGCT 569  
QY 402 CCTCTGGACATCAGCAGACATTTGGGGTCAAGACCGAGGCGCGGCTGTGCACTC 461  
Db 570 CCTTCCGACATCGACGAGTTCTTCAACATCGGCGACAGTAGCTCTGGGCTCATCCAGAC 629  
QY 462 AGTGTTCATCTGTAGCTTCAATGCTGCGCCCATCTTGGGCTACCTGGGCGACCGCTT 521  
Db 630 CGTGTTCATCTCCAGTTACATGTTGGCACTGTGTTGGCTACCTGGGTGACAGGTA 689  
QY 522 CAACAGAGAGTGAATCTCAGCTGGCGCAATTTTCTTCTGGTGGCGCGCTCACCTTCTCCAG 581  
Db 690 CAATCGAAGATATCTCATGTGGGGGCAATTCGCTTCTGTTCCCTGGTGACACTGGGGTC 749  
QY 582 CTCTTTCATTTCCAGCAGTACTTCTGGCTGTGCTCTGTCCTCGGGGGCTGGTGGGCT 641  
Db 750 ATCTTTCATCTCCCGAGAGCATTTCTGGCTGTCTCTCTGACCCCGGGGCTGGTGGGCT 809  
QY 642 CGGGAGGCGAGCTACTTCCACCATCGCCCACTATCATTTGGCGACCTCTTCCACCAAGAA 701  
Db 810 CGGGAGGCGAGTATTTCCACCATCGCCCACTCTCATTTGCCGACCTCTTTTGTGGCGGA 869  
QY 702 CACGGTACGCTCATGCTGTCTGCTTCTTCTTCTGCGCATTCACACTGGGAGTGGCTGG 761  
Db 870 CAGCGGAGCGGAGTGTCTGACATCTTCTTCTTCTTCTGCGGCTGGCTGGCTGG 929  
QY 762 CTACATTACTGCTCCAGCTGAAGCAGGAGCGGAGCTGGGCACTGGGCACTGGGCTGG 821  
Db 930 CTACATTGAGGCTCCAAAGTGAAGGATATGCTGGGAGATGCGCTGGGAGCTGGGCACTGGGCTGG 989  
QY 822 GTCCCTCTCTGGGCGATGATCACAGGAACATCATCTCTCATTTCTGGTCCCGGCACTAA 881

Db 990 GACACCGGCTAGAGTGGTGGCGCTTCTGCTGCTTCTGCTAGTGGCGGAGCCGCC 1049  
Qy 882 AAGGGGTATCCGACAGCTCGGGA---CCAGCTAAGCGCCGACCTCATGCTCCG 938  
Db 1050 AAGGGAGCGTGAGCGCACTCAGATTGCGACCCCTGAACCCACCTCGTGGTGGC 1109  
Qy 939 AGATATGAAGCCCTGTATTGAAACCGCAGTACGTCTTCTCTCCCTGGCCACGTGGC 998  
Db 1110 AGATCTAGGGCTCTGGCAAGAAATCCTAGTTTCTGCTGCTTCTCCCTGGCTTCACTGC 1169  
Qy 999 TGTCTCTTCCGACCGGGGCTTGGCATGTGATCCGCTCTACCTGACCCGCGCA 1058  
Db 1170 TGTGGCTTTGTACCGGCTCCCTGGCTGTGGGCTCCGCAATCTCTGCTGCTCCG 1229  
Qy 1059 AGTTGTGCA---GAAGACAGCAGAGACGTGCAACAGCCCGCTGTGGGCAAGGACAG 1115  
Db 1230 CGTGGCTTGGGAGACCCACCCCTGCTCCCGAGATCTCTGCTTCTCTGACAG 1289  
Qy 1116 CTTATCTTTGGGCCATCACTGCTTTTACGGGATTTCTGGGCTGTGTCAGCGGGCAGG 1175  
Db 1290 TCTCATCTTTGGACTCATCTGCTGACCGGAGTCTTGGGTGTGGGCTGGGTGGA 1349  
Qy 1176 AGCACCGCTGGTGGCTGAAAGACCCAGCGGGCGACCACTGTTGTGCTGGTGG 1235  
Db 1350 GATCAGCGCGGGCTCGGCCACTCCAAACCCCGGGCTGATCCCTGCTGTGTCACCTGG 1409  
Qy 1236 CATGCTGGGCTCTGCTATCTTCTGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCT 1295  
Db 1410 CTTCTGGGCTCTGACCTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1469  
Qy 1296 AGGAGCTATATCTGTTCTTCTGCGGGAGACCTGCTGTTTCTAACTGGGCGCATCAC 1355  
Db 1470 GGCACATTATTTTCTATCTTCTGAGAGACCTCTCTGTCATGACTGGGCTATCGT 1529  
Qy 1356 TGCAGACATCTCATGTAGTGTGTCATCCCAACCGGGCGGCACTGCTGCTGCTGCTGCA 1415  
Db 1530 GGGCGACATCTGCTGAGTGTGATCTTCTACCGACGCTTCCACCGCGGAGGCTTCCA 1589  
Qy 1416 GAGCTTCACTCTCCACCTGCTGGGGAGCGCGGGAGCCCTTACCTCATTTGGCTTTATCTC 1475  
Db 1590 GATGCTGCTGCTCCACCTGCTGGGTGATGCTGGAGGCGCTTACCTCATTTGGCTGATCTC 1649  
Qy 1476 AGACCTGATCCGCGAGACATAAGGACTCCCGCTCTGGAGTTCCTGAGCCTGGGCTA 1535  
Db 1650 TGACCGCTGGCGGAACTGGCCCCCTCTTCTTGTTCGAGTTCGGGCTCTGCAATT 1709  
Qy 1536 CGGCTCATGCTTGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1595  
Db 1710 CTGCTCATGCTGCTGCTGCTTGTGGGCACTGGGGGCGAGCTTCTGCGGACCGC 1769  
Qy 1596 GCTCTTCTGCTGAGCGACCGCGGCTGAGCAGCCTGGGG 1642  
Db 1770 CATCTTCAATTGAGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1816

RESULT 13  
AX481746  
LOCUS AX481746 2230 bp DNA linear PAT 16-AUG-2002  
DEFINITION Sequence 4 from Patent WO02055701.  
ACCESSION AX481746  
VERSION AX481746.1 GI:22316580  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1  
Curtis, R. A.  
Human sugar transporter proteins, potassium channel proteins,  
phospholipid transporter proteins and methods of use thereof  
Patent: WO 02055701-A 4 18-JUL-2002;  
Millennium Pharmaceuticals, Inc. (US)

FEATURES  
source  
CDS  
Location/Qualifiers  
1..2230  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
376..1962  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD44420.1"  
/db\_xref="GI:22316581"  
/translation="MAGSDTAPFLSQADDDPDDPVPCTPGLPGSTGNPKSEPEVPDQ  
EGQRIITGLSFRSALIVAVLCYINLNTMDRFTVAGVLDPIEDIFNIGDSSGLIQT  
VFSSVNLAPVFGYLDNRKYLKMGGIAFMSLVLTGSSPIGEPHFWLLLLTGLV  
GVEASSTIATVLLADLVADORSRLSIFYPALPVNGSLGYTAGSKVDMAGDWH  
ALRVTGLVAVLLPLVVRPGRGALVHSDLPPLNPTSMWADLALARNPSVLS  
SLGTAFAVFTGLSALWAPFLARSVLGETPCLPGDCSSDSLI.FGLITCLTGV  
LVLGLVEISRRLRHNPADPLVATGLGSLAPPLFLSLACAGRSIVATYIFPIGE  
TLGSMNVAIVADILLIYVPIPTRRSTAEAFQIVLSHLIGDAGSPYLIGLISLRNRNP  
PSFLSEFRALQFSLMLCAFVAGLGAFLGTAIFIEADRRRAQLHVQGLLHAGSTDD  
RIVVPQRGRSTRVPAVSLI"  
ORIGIN  
Query Match 27.2%; Score 563.8; DB 6; Length 2230;  
Best Local Similarity 64.3%; Pred. No. 3e-72;  
Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;  
Qy 282 GGGCCCGCGCTCAGCAGCCCAACCGCGCAGCTTGGCGCGCGGGGGGCGCGC 341  
Db 501 GGACAGAGGGCTGACGCGCATCACCGGCTGCTCTCCCGCGCGTTCGGCTCTCATAGT 560  
Qy 342 CGCCATCTCAGCTTGGGCAACGCTGCTCAACTACTGACAGGTACACCGTGGCAGGCGT 401  
Db 561 GCGGCTGCTGCTACATCAATCTCTGAACTACATGACCGCTTACCGTGGCTGGGT 620  
Qy 402 CTTTCTGGAATCCAGAGCAGCTTTGGGGTCAAGGACCGAGCGCCGCGCTGCTGCAATC 461  
Db 621 CTTTCCGACATCGAGCAGTCTTCAACATCGGGGACAGTAGTCTTGGGCTCATCCAGC 680  
Qy 462 AGTTTCTATCTGAGCTTCTATGCTGGCTGCGCCCATCTTGGCTTACCTGGCGCAGCGCTT 521  
Db 681 CGTGTTCATCTCAGTTACATGTTGGCACCTGTGTTGGCTTACCTGGGTGACAGGTA 740  
Qy 522 CAACAGAGAGTGAATCTCAGCTCGCGCATTTTCTTCTGCTGCGCGCTTCTTCTCCAG 581  
Db 741 CAATCGAGAGTATCTCATGTCGCGGGGCAATGCTTCTGCTTCTGCTGCTGCTGCTGCT 800  
Qy 582 CTCCTTCATTCGCCAGAGTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 641  
Db 801 ATCTTTCATTCGCCAGAGCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 860  
Qy 642 CGGGGAGCGCAGTACTCCACCATCGCCCGCATATCATTTGCGGAGCTTCTTCCACCAAGNA 701  
Db 861 CGGGGAGCGCAGTTATTTCCACCATCGCGCCCATCTCATTTGCGGAGCTTCTTGTGGCGGA 920  
Qy 702 CACGCTTACCTCATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 761  
Db 921 CCAGCGAGCGCGATGCTCAGCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 980  
Qy 762 CTACATTAATTCAGCTGTAAGAGCGAGCGCGGAGCTGCGCATCTGCGCATCTGCGGT 821  
Db 981 CTACATTTGAGGCTCCAAAGTGAAGATATGGCTGGAGACTGCGCATCTGCGGTCTGAGGGT 1040  
Qy 822 GTCCCTGCTCTGGGATGATCAGGAGACATCTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 881  
Db 1041 GACACCGGCTTAGGAGTGGTGGCGCTTCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1100  
Qy 882 AAGGGCTCATGCGGACAGCTCGGGA---CCAGCTCAAGCGCGCGGACCTCATGGCTCCG 938  
Db 1101 AAGGGAGCGGTGAGGCGCACTCAGATTGGCCACCCCTGAACCCACCTCTGTTGGTGGG 1160  
Qy 939 AGATATGAAGCGCTGATTCGAAACCGAGCTACGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 998  
Db 1161 AGATCTGAGGCTCTGCGAAGAAATCCTAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1220

QY	999	TGTCCTCTTGCCACGCGGGCCCTGGCATGTGGATCCGCTACTCTGACACCGCGCCA	1058
Db	1221	TGTGGCCTTTGTACGCGGCTCCCTGGCTGTGTGGCTCCGGCATTTCTGCTCGCTCCCG	1280
QY	1059	AGTTGTGCA--GAGACAGCAGACGCTGCAACAGCCGCGCTGTGGGGCCAAGACAG	1115
Db	1281	CGTGTCTTTGGGAGACCCACCTCGCTTCCCGGAGACTCTGCTCTTCTCTCTGACAG	1340
QY	1116	CTCATCTTTGGGGCATCACTGCTTTACGGGATTTCTGGCGCTGTGTCACGGGGCAGG	1175
Db	1341	TCTCATCTTTGAGACTCATCACTGCTCAGCGGAGTCTCGGTGTGGCCTGGGTGTGGA	1400
QY	1176	AGCCACGCGCTGTGTCCGCTGAAGACCCAGCGGGCGACCCACTGTGTGTGTCGGTGG	1235
Db	1401	GATCAGCGCGCGCTCCGCCACTCCAAACCCCGGGCTGATCCCTGTGTGTGCCACTGG	1460
QY	1236	CATGCTGGGCTCTGCCATCTTCATCTGCTGATCTTGTGGCTGCCAAGACGACATCGT	1295
Db	1461	CTCTCTGGGCTCTGACCCCTTCTCTTCCCTGTCCCTTGGCTGGCGCCGCTGTATGCA	1520
QY	1296	AGGAGCCTATATCTGTATCTTCTGGGGAGACGCTGTGTCTTAACTGGGCGCATCAC	1355
Db	1521	GGCCACTTATATTTTTCATCTTCAITGGAGAGACCTCTCTGTCCATGAATCGGCCATCGT	1580
QY	1356	TGCAGACATCTCATGTACGTGTTCATCCACGCGGGCGGCACCTGCCGTGGCTTGCA	1415
Db	1591	GGCGACATTCGTGTGTACGTGTGATCCCTACCGACGCTCCACGCGGAGGCTTCCA	1640
QY	1416	GAGCTTCACTCCACCTGCTGGGGGACGCGGGAGCCCTACCTCATTTGGCTTTATCTC	1475
Db	1641	GATCGTGTGTGCCACCTGCTGGGTGTGTGGGAGCCCTAGCTCATTTGGCCTGATCTC	1700
QY	1476	AGACCTGATCGCCAGACACTAAGGACTCCCGCTCTGGGAGTTCCTGAGCTGGGCTA	1535
Db	1701	TGACCGCTGGCGCGGAACGGCCCCCTCTCTTGTCCGAGTTCCGGGCTCTGAGTT	1760
QY	1536	CGCGCTCATGCTCGCCCTTTGTCGTGTGTCTGGGGGCAATGTTCTTCTCGCCACTGC	1595
Db	1761	CTCGCTCATGCTCTGCGCGTTTGTGTGGGCACTGGGCGGGCGAGCCCTTCTGGGACCGC	1820
QY	1596	GCTCTTCTTGTCAGACCGCGCCACGGGCTGACGACCACTGGGGG	1642
Db	1821	CATCTCATTTGAGCGACCGCGCGGGGACAGCTGTCAGCTTCAGG	1865

**RESULT 14**

LOCUS	BD191275	2201 bp	DNA	linear	PAT 17-JUL-2003
DEFINITION	186 human secreted proteins.				
ACCESSION	BD191275				
VERSION	BD191275.1	GI:33001014			
KEYWORDS	JP 2002510192-A/239.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 2201)				
AUTHORS	Ruben, S.M., Rosen, C.A., Fischer, C.L., Soppet, D.R., Carter, K.C., Bednarek, D.P., Endress, G.A., Yu, G.L., N.J., Feng, P., Young, P.E., Greene, J.M., Ferrie, A.M., Duan, R., Hu, J.S., Florence, K.A., Olsen, H.S., Ebner, R., Brewer, L.A., Moore, P.A., Shi, Y., Lafleur, D.W., Li, Y., Zeng, Z. and Kyaw, H.				
TITLE	186 human secreted proteins				
JOURNAL	Patent: JP 2002510192-A 239 02-APR-2002;				
COMMENT	HUMAN GENOME SCIENCES INC PN JP 2002510192-A/239 PD 02-APR-2002 PF 06-MAR-1998 JP 1998538883 PR 07-MAR-1997 US 60/040162, 07-MAR-1997 US 60/040333 PR 07-MAR-1997 US 60/038621, 07-MAR-1997 US 60/040161 PR 07-MAR-1997 US 60/040826, 07-MAR-1997 US 60/040334 PR 07-MAR-1997 US 60/040336, 07-MAR-1997 US 60/040163 PR 11-APR-1997 US 60/043580, 11-APR-1997 US 60/043568 PI				



Qy	1299	AGCTATATCTGTATCTTCGTGGGAGAGCGCTGCTGTTTCTTAAC	TGGGCCATCACTGC	1358
Db	1435	CACTTATATTTTATCTTTCATTGGAGAGACCTCTCTGTCCATGA	AACTGGGCCATCGTGGC	1494
Qy	1359	AGACATCCTCATGTACGTGGTCA	TCCCAACGCGGCGCCTGCGCTGGCCTTG	CAGAG 1418
Db	1495	CGACATTCGTCTGTACGTGGTATCCCTACCGAGCTCCACCGCGAGGCCTTCCAGAT		1554
Qy	1419	CTTCACCTCCCACTGCTGGGGAGCCCGGGAGCCCTACCTCATTTGGCTTTATCTCAGA		1478
Db	1555	CGTGTGTCCTCCACCTGCTGGGTGATGCTGGGAGCCCTACCTCATTTGGCCTGATCTCTGA		1614
Qy	1479	CCTGATCCGCCAGAGCACTAAGGACTTCCCGCTCTGGGAGTTCCCTGAGCCTGGGCTACGC		1538
Db	1615	CGGCCTGGCGCGAACTGGCCCCCTCTCTTGTCCGAGTCCGGGCTCTGCAGTTCTC		1674
Qy	1539	GCTCATGCTCTGCCCTTTGCTGCTGGTCTCTGGGGGCATGTTCTTCTCGCCACTGGGCT		1598
Db	1675	GCTCATGCTCTGGCGTTTGTGGGGCACTGGGGCGCACT-TTCCTGGGCACCGCCAT		1733
Qy	1599	CTTCTTCTGTCAGCGACCGCGCCAGGGCTGAGCAGCACTGGGGG		1642
Db	1734	CTTCATTGAGGCCGACCGCGGGGCACAGCTGCAGTGCAGG		1777

Search completed: October 23, 2004, 11:01:11  
Job time : 8813 secs





PR 12-APR-2001; 2001US-0283443P.  
PR 23-APR-2001; 2001US-0285754P.  
PR 24-APR-2001; 2001US-0286096P.  
PR 03-MAY-2001; 2001US-0286353P.  
PR 17-MAY-2001; 2001US-0291703P.  
PR 31-MAY-2001; 2001US-0294834P.  
PR 20-JUN-2001; 2001US-0295695P.  
PR 21-JUN-2001; 2001US-0298845P.  
PR 05-JUL-2001; 2001US-0303242P.  
PR 13-AUG-2001; 2001US-0311981P.  
PR 16-AUG-2001; 2001US-0312858P.  
PR 17-AUG-2001; 2001US-0313280P.  
PR 29-AUG-2001; 2001US-0315614P.  
PR 17-SEP-2001; 2001US-0322818P.  
PR 25-FEB-2002; 2002US-00322818.  
XX (CURA-) CURAGEN CORP.  
XX  
PI Alsobrook JP, Anderson DW, Ballinger RA, Boldog FI, Burgess CE;  
PI Casman SJ, Ellerman KE, Gangolli RA, Gerlach VI, Gilbert JA;  
PI Gorman L, Guo X, Gusev VY, Kekuda R, Li L, Liu X, Malyankar UM;  
PI Miller CE, Millet I, Padigaru M, Patturajan M, Pena CE, Peyman JA;  
PI Rastelli L, Shenoy SG, Shimkets RA, Smithson G, Spytek KA, Stone DJ;  
PI Taupier RU, Tchernev VT, Vernet CAM, Zerhusen BD;  
XX  
DR WPI; 2002-698672/75.  
DR P-PSDB; ADH48764.  
XX  
PT New NOVX polypeptides or polynucleotides, useful for preventing or  
PT treating disorders or syndromes e.g., atherosclerosis, hypertension,  
PT obesity or cancer.  
XX  
PS Claim 8; Page 129-130; 923pp; English.  
XX  
CC The present invention relates to novel human NOVX proteins, where X is  
CC any number from 1 to 91 and their coding sequences. The proteins and  
CC coding sequences are useful for preventing or treating disorders or  
CC syndromes e.g. atherosclerosis, hypertension, obesity or cancer. NOV19 is  
CC a spinster-like protein and its coding sequence maps to chromosome 17.  
XX  
SQ Sequence 2071 BP; 341 A; 681 C; 685 G; 364 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 2071; DB 6; Length 2071;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2071; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CCCCCCGCCCGCCGATCCGGGCGGCATGATGCTGCTGGAATGCGCCCTCGCGGGCGG 60  
DB 1 CCCCCCGCCCGCCGATCCGGGCGGCATGATGCTGGAATGCGCCCTCGCGGGCGG 60  
  
QY 61 GCGGCGCGGAGGAGGAGGCGGACGCGGAGCGCGCGCGCGCGCGCGGGGGCGGAG 120  
DB 61 GCGGCGCGGAGGAGGAGGAGGCGGACGCGGAGCGCGCGCGCGCGCGGGGGCGGAG 120  
  
QY 121 CGAGGGGCTGGCGGTAGCGGTTGCTGCGGGGCGCGGGGCGCGGGGCGCGGCTGCG 180  
DB 121 CGAGGGGCTGGCGGTAGCGGTTGCTGCGGGGCGCGGGGCGCGGGGCGCGGCTGCG 180  
  
QY 181 GCGCGGGGCGATGAGGTGCGAGACGCTGTGCGGCGAGCGGTAAAGCGGCGCCCGACCGGACCC 240  
DB 181 GCGCGGGGCGATGAGGTGCGAGACGCTGTGCGGCGAGCGGTAAAGCGGCGCCCGACCGGACCC 240  
  
QY 241 CCGGCAACCCCGGACCCCGCGGTGCGAGCTACTGCAAGGGCGCGCGGCTCAGCAG 300  
DB 241 CCGGCAACCCCGGACCCCGCGGTGCGAGCTACTGCAAGGGCGCGCGGCTCAGCAG 300  
  
QY 301 CCAAAACCGGCGAGCTTGGGCGCGCGGGGCGGAGCGCGCCATCTCAGCTTGGGC 360  
DB 301 CCAAAACCGGCGAGCTTGGGCGCGCGGGGCGGAGCGCGCCATCTCAGCTTGGGC 360  
  
QY 361 AACGTGCTCAACTCTGAGACAGGTACACCGTGGGAGCGGTCTTCTTGACATCCAGCAG 420  
DB 361 AACGTGCTCAACTCTGAGACAGGTACACCGTGGGAGCGGTCTTCTTGACATCCAGCAG 420

QY 421 CACTTTGGGGTCAAGGACCGAGCGCGGCTGTGCTGAGTCACTGTTTCACTGTTAGCTTC 480  
DB 421 CACTTTGGGGTCAAGGACCGAGCGCGGCTGTGCTGAGTCACTGTTTCACTGTTAGCTTC 480  
  
QY 481 ATGGTGGCTGCCCCCATCTTCGGCTACCTGGGCGACCGCTTCAACAGGAAGGTGATTCTC 540  
DB 481 ATGGTGGCTGCCCCCATCTTCGGCTACCTGGGCGACCGCTTCAACAGGAAGGTGATTCTC 540  
  
QY 541 AGCTGGGCAATTTCTTCTGCTGGCGCTACCTTCTCCAGCTCTTCAATTTCCCGCAGAG 600  
DB 541 AGCTGGGCAATTTCTTCTGCTGGCGCTACCTTCTCCAGCTCTTCAATTTCCCGCAGAG 600  
  
QY 601 TACTTCTGGCTGCTGCTCTGCTCCGGGGGCTGGTGGGCACTCGGGAGGCGCAGCTACTCC 660  
DB 601 TACTTCTGGCTGCTGCTCTGCTCCGGGGGCTGGTGGGCACTCGGGAGGCGCAGCTACTCC 660  
  
QY 661 ACCATCGGCGCCCACTATCATTTGGGACCTTTCACCAAGAACACCGGTACGCTCATGCTG 720  
DB 661 ACCATCGGCGCCCACTATCATTTGGGACCTTTCACCAAGAACACCGGTACGCTCATGCTG 720  
  
QY 721 TCCGCTCTTCTACTTCCGCATCCCACTGGGCGAGTGGCTGGGCTACATTTACTGGCTCCAGC 780  
DB 721 TCCGCTCTTCTACTTCCGCATCCCACTGGGCGAGTGGCTGGGCTACATTTACTGGCTCCAGC 780  
  
QY 781 GTGAAGCAGGCGAGCGGAGACTGGGCACTGGGCACTGGGGTGTCCCTGTCTCTGGGCAATG 840  
DB 781 GTGAAGCAGGCGAGCGGAGACTGGGCACTGGGCACTGGGGTGTCCCTGTCTCTGGGCAATG 840  
  
QY 841 ATCAGAGAACACATCATCTCTATTTCTGTCGCCACCACTAAAGGGGTCAATGCCGACCGAG 900  
DB 841 ATCAGAGAACACATCATCTCTATTTCTGTCGCCACCACTAAAGGGGTCAATGCCGACCGAG 900  
  
QY 901 CTCGGGACCAAGCTCAAGCGCCCGGACCTCATGGCTCCGAGATATGAAGGCCCTCATTTCCA 960  
DB 901 CTCGGGACCAAGCTCAAGCGCCCGGACCTCATGGCTCCGAGATATGAAGGCCCTCATTTCCA 960  
  
QY 961 AACCGCAGTACGCTCTCTCTCTGTCGCCACGCTGGCTGTCTCTTCCGACCGGGGGCC 1020  
DB 961 AACCGCAGTACGCTCTCTCTCTGTCGCCACGCTGGCTGTCTCTTCCGACCGGGGGCC 1020  
  
QY 1021 CTGGGCAATGTGGATCCCGCTCTTACCTGCACCGCGCCCAAGTTGTGAGAGACAGCAGAG 1080  
DB 1021 CTGGGCAATGTGGATCCCGCTCTTACCTGCACCGCGCCCAAGTTGTGAGAGACAGCAGAG 1080  
  
QY 1081 ACGTGCAACAGCGCCGCTGTGGGCGCAAGGACAGCTCATCTTTTGGGGCGCATCACCTGC 1140  
DB 1081 ACGTGCAACAGCGCCGCTGTGGGCGCAAGGACAGCTCATCTTTTGGGGCGCATCACCTGC 1140  
  
QY 1141 TTTACGGGATTTCTGGGCGTGGTCAAGGGGCGAGGAGCCACGCTGGTGGGGCGCTGAAG 1200  
DB 1141 TTTACGGGATTTCTGGGCGTGGTCAAGGGGCGAGGAGCCACGCTGGTGGGGCGCTGAAG 1200  
  
QY 1201 ACCCAGCGGCGGACCCCACTGGTGTGCGTGGGCAATGCTGGGCTCTGCACTCTTCATC 1260  
DB 1201 ACCCAGCGGCGGACCCCACTGGTGTGCGTGGGCAATGCTGGGCTCTGCACTCTTCATC 1260  
  
QY 1261 TGCTGATCTTCTGGTGGTGGCAAGAGGAGCATCGTAGGACCTATATCTGTATCTTCGTC 1320  
DB 1261 TGCTGATCTTCTGGTGGTGGCAAGAGGAGCATCGTAGGACCTATATCTGTATCTTCGTC 1320  
  
QY 1321 GGGGAGACGCTGTGTTTTCTAACTGGGCGCATCACTGCGAGACATCTCATGTACGTGGTC 1380  
DB 1321 GGGGAGACGCTGTGTTTTCTAACTGGGCGCATCACTGCGAGACATCTCATGTACGTGGTC 1380  
  
QY 1381 ATCCCAACGGGCGGCGGCACTGGCTGGGCTTTCAGAGCTTCACTCCCACTGCTGGGG 1440  
DB 1381 ATCCCAACGGGCGGCGGCACTGGCTGGGCTTTCAGAGCTTCACTCCCACTGCTGGGG 1440  
  
QY 1441 GAGCGGGAGGCGCCCTACCTCATTTGGCTTTATCTCAGACCTGATCGGCGAGACACTAAG 1500  
DB 1441 GAGCGGGAGGCGCCCTACCTCATTTGGCTTTATCTCAGACCTGATCGGCGAGACACTAAG 1500



Db 241 GCGGGGGGNAAGCGCGGCCATCCTCAGCTTGGGCAACGTGCTCAACTACCTGGACAGGT 300  
Qy 386 ACACCGTGGCAGCGCTCTTCTGGACATCCAGCAGCACTTTGGGGTCAAGGACCGAGGCG 445  
Db 301 ACACCGTGGCAGCGCTCTTCTGGACATCCAGCAGCACTTTGGGGTCAAGGACCGAGGCG 360  
Qy 446 CCGGCTGTGAGTCAAGTGTTCATCTGTAGCTTTCATGTGGTGGTGGCCCATCTTCGGGT 505  
Db 361 CCGGCTGTGAGTCAAGTGTTCATCTGTAGCTTTCATGTGGTGGTGGCCCATCTTCGGGT 420  
Qy 506 ACCTGGGCGACCGCTTCAACAGGAAGTATCTCAGCTGCGGCATTTTCTTCGTGTCG 565  
Db 421 ACCTGGGCGACCGCTTCAACAGGAAGTATCTCAGCTGCGGCATTTTCTTCGTGTCG 480  
Qy 566 CCGTCACTTCTCCAGTCTCTTCATTTCCCGCAGTACTTCTGGCTGTGCTGCTCCTGCC 625  
Db 481 CCGTCACTTCTCCAGTCTCTTCATTTCCCGCAGTACTTCTGGCTGTGCTGCTCCTGCC 540  
Qy 626 GGGGGCTGTGGGATCGGGGAGCCAGCTACTCCACCATCGCCCCCATCATCATTTGGCG 585  
Db 541 GGGGGCTGTGGGATCGGGGAGCCAGCTACTCCACCATCGCCCCCATCATCATTTGGCG 600  
Qy 686 ACCTCTTCAACAAAGAACCGCTACGCTCATGCTGCTCGTCTTCTACTTCGCGCATCCAC 745  
Db 601 ACCTCTTCAACAAAGAACCGCTACGCTCATGCTGCTCGTCTTCTACTTTCGCCATCCAC 560  
Qy 746 TGGGCAGTGGCTTGGGCTACATTTACTGGCTTCCAGCTGAAGCAGGAGCGGAGACTGGC 805  
Db 661 TGGGCAGTGGCTTGGGCTACATTTACTGGCTTCCAGCTGAAGCAGGAGCGGAGACTGGC 720  
Qy 806 ACTGGGCATTTGGGCTGCTCCCTGCTCGGGCATGATCACGGAACACTCATCTCTATTC 865  
Db 721 ACTGGGCATTTGGGCTGCTCCCTGCTCGGGCATGATCACGGAACACTCATCTCTATTC 780  
Qy 866 TGTCTCCAGCACTAAAGGGGTTCATGCCGACAGCTCGGGGACACAGCTCAAGGCCCGGA 925  
Db 781 TGTCTCCAGCACTAAAGGGGTTCATGCCGACAGCTCGGGGACACAGCTCAAGGCCCGGA 840  
Qy 926 CCTCATGGTCCGAGATATGAAGGCGCTGATTCGAAACCGCAGCTACGTCTTCTCTCCC 985  
Db 841 CCTCATGGTCCGAGATATGAAGGCGCTGATTCGAAACCGCAGCTACGTCTTCTCTCCC 900  
Qy 986 TGGCCAGTGGCTGCTCTCTCGCCAGCGGGGCGCTGGGCACTGGATCCGCTCTACC 1045  
Db 901 TGGCCAGTGGCTGCTCTCTCGCCAGCGGGGCGCTGGGCACTGGATCCGCTCTACC 960  
Qy 1046 TGACCGCGCCCAAGTTGTGCAAGACAGCAGACGTGCAACAGCGCGCCCTGTGGGG 1105  
Db 961 TGACCGCGCCCAAGTTGTGCAAGACAGCAGACGTGCAACAGCGCGCCCTGTGGGG 1020  
Qy 1106 CCAAGGACAGCTCATCTTTGGGGCCATCATCTGCTTTTACGGATTTCTGGGGTGGTCA 1165  
Db 1021 CCAAGGACAGCTCATCTTTGGGGCCATCATCTGCTTTTACGGATTTCTGGGGTGGTCA 1080  
Qy 1166 CCGGGGAGGAGCAGCGCTGTGCGCTGAAGACCCAGCGGGCGGACCCACTGGTCT 1225  
Db 1081 CCGGGGAGGAGCAGCGCTGTGCGCTGAAGACCCAGCGGGCGGACCCACTGGTGT 1140  
Qy 1226 GTCCGCTGGGCACTGCTGGGCTGCCATCTTCATCTGCTGATCTTCTGGGCTGCCAAGA 1285  
Db 1141 GTCCGCTGGGCACTGCTGGGCTGCCATCTTCATCTGCTGATCTTCTGGGCTGCCAAGA 1200  
Qy 1286 GCAGCATGTAGAGGCTATATCTGTATCTTCTGGGAGAGCGTGTCTTCTAACT 1345  
Db 1201 GCAGCATGTAGAGGCTATATCTGTATCTTCTGGGAGAGCGTGTCTTCTAACT 1260  
Qy 1346 GGGCCATCATCTGCAGACATCCTCATGTGTGTTCATCCCGCGGGCGCCACTGCGG 1405  
Db 1261 GGGCCATCATCTGCAGACATCCTCATGTGTGTTCATCCCGCGGGCGCCACTGCGG 1320  
Qy 1406 TGGCTTGCAGAGCTTTCACCTCCCATCTGCTGGGAGCGCGGAGCCCTACCTCATTTG 1465  
Db 1321 TGGCTTGCAGAGCTTTCACCTCCCATCTGCTGGGAGCGCGGAGCCCTACCTCATTTG 1380

Qy 1466 GCTTATCTCAGACCTGATCCGCCAGAGCACTAAGGACTCCCCGCTCTGGGAGTTCCTGA 1525  
Db 1381 GCTTATCTCAGACCTGATCCGCCAGAGCACTAAGGACTCCCCGCTCTGGGAGTTCCTGA 1440  
Qy 1526 GCTCGGGCTACGCGCTCATGCTCTGCCCTTTGCTGCTGCTGGCGGCATGTTCTTCC 1585  
Db 1441 GCTCGGGCTACGCGCTCATGCTCTGCCCTTTGCTGCTGCTGGCGGCATGTTCTTCC 1500  
Qy 1586 TCGCCACTGCGCTCTTCTTCTGTCAGCGACCGCGCCAGGGCTGAGCAGCACCTG 1638  
Db 1501 TCGCCACTGCGCTCTTCTTCTGTCAGCGACCGCGCCAGGGCTGAGCAGCACCTG 1553  
RESULT 3  
ADBE07415  
ID ADE07415 standard; DNA; 1250 BP.  
XX AC ADE07415;  
XX DT 29-JAN-2004 (first entry)  
XX DE Novel coding sequence (useful for identifying genetic disorders) #481.  
XX KW novel gene; novel protein; tissue marker; molecular weight marker;  
XX KW chromosome marker; genetic disorder; gene; ds.  
XX OS Unidentified.  
XX PN WO2003054152-A2.  
XX PD 03-JUL-2003.  
XX PF 10-DEC-2002; 2002WO-US039555.  
XX PR 10-DEC-2001; 2001US-0339739P.  
XX PR 11-DEC-2001; 2001US-0339453P.  
XX PR 14-MAR-2002; 2002US-0365091P.  
XX PR 14-MAR-2002; 2002US-0365384P.  
XX PR 12-APR-2002; 2002US-0372381P.  
XX PR 22-APR-2002; 2002US-0372615P.  
XX PR 24-APR-2002; 2002US-0376045P.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;  
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;  
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;  
XX WPI; 2003-569235/53.  
DR P-PSDB; ADE08326.  
XX New polynucleotides, useful for expressing recombinant proteins for  
PT analysis, characterization or therapeutic use, or as markers for tissues  
PT in which the corresponding protein is preferentially expressed.  
XX Claim 1; SEQ ID NO 481; 1177pp; English.  
XX The invention comprises the amino acid and coding sequences of novel  
CC proteins. The DNA and protein sequences of the invention are useful as:  
CC markers for tissues in which the corresponding protein is preferentially  
CC expressed; as molecular weight markers on gels; as chromosome markers or  
CC tags; to identify chromosomes or to map related gene positions; and to  
CC compare with endogenous DNA sequences in patients to identify potential  
CC genetic disorders. The present DNA sequence represents a gene of the  
CC invention.  
XX  
SQ Sequence 1250 BP; 202 A; 407 C; 419 G; 222 T; 0 U; 0 Other;  
Query Match 45.1%; Score 933.4; DB 10; Length 1250;  
Best Local Similarity 99.9%; Pred. No. 2.3e-161;  
Matches 934; Conservative. 0; Mismatches 1; Indels 0; Gaps 0;



Db	121	TTGTGAGAGAAGCAGCAGAGACGTGCAACAGCCGCGCTGTGGGCCAAGGACAGCCTCAT	180
Qy	1122	CTTTGGGGCCATCACTCTGTTTACGGGATTTCTGGCGGTGTGTCA CGGGGGCAGAGCCAC	1181
Db	181	CTTTGGGGCCATCACTCTGTTTACGGGATTTCTGGCGGTGTGTCA CGGGGGCAGAGCCAC	240
Qy	1182	GGCGTGTGCGCCCTGAAGACCCAGCGGGCCGACCCACTGGTGTGTGCGGTGGGCATGCT	1241
Db	241	GGCGTGTGCGCCCTGAAGACCCAGCGGGCCGACCCACTGGTGTGTGCGGTGGGCATGCT	300
Qy	1242	GGGCTCTGCCATCTTCATCTGCCGTGATCTCGTGGCTGCCAAGAGCAGCATCTGTAGGAGC	1301
Db	301	GGGCTCTGCCATCTTCATCTGCCGTGATCTCGTGGCTGCCAAGAGCAGCATCTGTAGGAGC	360
Qy	1302	CTATATCTGTATCTTCGTGCGGGAGACGCTGCTGCTTTTCTAACTGGGCGCATCACTGCAGA	1361
Db	361	CTATATCTGTATCTTCGTGCGGGAGACGCTGCTGCTTTTCTAACTGGGCGCATCACTGCAGA	420
Qy	1362	CATCCTCATGTACGTGTGCATCCCAACGCGGGCGGCCACTGCCGTGGCTTCGAGAGCTT	1421
Db	421	CATCCTCATGTACGTGTGCATCCCAACGCGGGCGGCCACTGCCGTGGCTTCGAGAGCTT	480
Qy	1422	CACCTCCACCTGTCTGGGGACGCGGGAGCGCCCTACCTCATTTGGCTTTTATCTCAGACCT	1481
Db	481	CACCTCCACCTGTCTGGGGACGCGGGAGCGCCCTACCTCATTTGGCTTTTATCTCAGACCT	540
Qy	1482	GATCCGCGAGCACTAAGACATCCC CGCTCTGGAGTTCCCTGAGCGTGGGCTACGCGCT	1541
Db	541	GATCCGCGAGCACTAAGACATCCC CGCTCTGGAGTTCCCTGAGCGTGGGCTACGCGCT	600
Qy	1542	CATGCTCTGCCCTTTCGTCTGTGTCTTGGCGGCATGTTCTTCTCGCCACTGCGCTCTT	1601
Db	601	CATGCTCTGCCCTTTCGTCTGTGTCTTGGCGGCATGTTCTTCTCGCCACTGCTCTT	660
Qy	1602	CTTTCGTACGACGCGCGCCAGGCGTGTGAGCAGCACCTG	1638
Db	661	CTTTCGTACGACGCGCGCCAGGCGTGTGAGCAGCACCTG	697

RESULT 5	
AAI59329	
ID	AAI59329 standard; cDNA; 1725 Bp.
XX	
XX	AAI59329;
XX	
XX	
XX	22-OCT-2001 (first entry)
XX	
XX	Human polynucleotide SEQ ID NO 1532.
XX	
XX	Human; neutropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX	leukaemia; ss.
XX	
XX	Homo sapiens.
XX	
XX	WO200153312-A1.
XX	
XX	26-JUL-2001.
PD	
XX	
XX	26-DEC-2000; 2000WO-US034263.
XX	
XX	23-DEC-1999; 99US-00471275.
PPR	21-JAN-2000; 2000US-00488725.
PPR	25-APR-2000; 2000US-00552317.
PPR	20-JUN-2000; 2000US-00598042.
PPR	19-JUL-2000; 2000US-00620312.
PPR	03-AUG-2000; 2000US-00653450.
PPR	14-SEP-2000; 2000US-00662191.
PPR	19-OCT-2000; 2000US-00693036.

PR	29-NOV-2000; 2000US-00727344.
XX	(HYSE-) HYSEQ INC.
XX	
XX	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI	Zhou F, Goodrich R, Drmanac RT;
XX	WPI; 2001-442253/47.
DR	P-PSDB; AAM40173.
XX	
XX	Novel nucleic acids and polypeptides, useful for treating disorders such
PT	as central nervous system injuries.
PT	
XX	Claim 1; SEQ ID NO 1532; 10078pp; English.
PS	
XX	
XX	The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC	encoded polypeptides (AAM38642-AAW42213) with neurotropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilisation of the activities such as: Immune system suppression,
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders. Note: The sequence data for this patent did not form
CC	part of the printed specification

Query Match	29.2%;	Score 605.4;	DB 4;	Length 1725;
Best Local Similarity	99.0%;	Pred. No. 2.2e-101;		
Matches 609;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
QY	1024	GGCATGTGGATCCGCTCTACCTGCA CGCGCCCAAGTTGTGCAGAGACAGACGACG 108		
Db	15	GGCATGTGGATCCCGCTCTACTCGACCGCGCCCTTGTGTGCATATAGACAGCAGACG 74		
QY	1084	TGCAACAGCCGCGCTCTGGGGCCAAAGACAGACCTCATCTTTGGGGCCATCACTGCTTT 1143		
Db	75	TGCAACAGCCGCGCTGTGGGCCAAGACAGACCTCATCTTTGGGGCCATCACTGCTTT 134		
QY	1144	ACGGGATTTCTGGCGTGGTCA CGGGGGCAGGAGCCACGCGCTGGTGC CGCTCAAGACC 1203		
Db	135	ACGGGATTTCTGGCGTGGACA CGGGGGCAGGAGCCACGCGTGGTGC CGCTCAAGACC 194		
QY	1204	CAGGGGCCGACCCACTGGTGTGTCGCGTGGGCATGCTGGGCTCTGCATCTTCACTGC 1263		
Db	195	CAGGGGCCGACCCACTGGTGTGTCGCGTGGGCATGCTGGGCTCTGCATCTTCACTGC 254		
QY	1264	CTGATCTTCGTGGTGCAGAGACAGCATCGTAGGAGCCATATCTGTATCTTCGTGGG 1323		
Db	255	CTGATCTTCGTGGTGCAGAGACAGCATCGTAGGAGCCATATCTGTATCTTCGTGGG 314		
QY	1324	GAGACGCTGCTGTTTCTTAACTGGGGCCATCACTGCGAGACATCCCTCATGTACGTGGTCATC 1383		
Db	315	GAGACGCTGCTGTTTCTTAACTGGGGCCATCACTGCGAGACATCCCTCATGTACGTGGTCATC 374		
QY	1384	CCACCGGGCGGCGCACTGCCGTGGCTTGCAGAGCTTCACTCCACCTGCTGGGGGAC 1443		
Db	375	CCACCGGGCGGCGCACTGCCGTGGCTTGCAGAGCTTCACTCCACCTGCTGGGGGAC 434		
QY	1444	GCGGGAGCCCTTACCCTCATTTGGCTTTATCTCAGACCTGATCCGCGCAGACACTAAGGAC 1503		
Db	435	GCGGGAGCCCTTACCCTCATTTGGCTTTATCTCAGACCTGATCCGCGCAGACACTAAGGAC 494		
QY	1504	TCCCGCTCTGGAGTTCGACCTCGGGCTACGGCGCTCATGCTCGCCCTTTTCGTCGTG 1563		
Db	495	TCCCGCTCTGGAGTTCGACCTCGGGCTACGGCGCTCATGCTCGCCCTTTTCGTCGTG 554		

Qy 1564 GTCCTGGCGGCATGTTCTTCTCGCCACATCGCTCTTCTTCGTGACGACCGCGCCAGG 1623  
Db |||||  
Qy 555 GTCCTGGCGGCATGTTCTTCTCGCCACATCGCTCTTCTTCGTGACGACCGCGCCAGG 614  
Db |||||  
Qy 1624 GCTGAGCAGCACCTG 1638  
Db |||||  
Qy 615 GCTGAGCAGCAGGTG 629  
Db |||||

## RESULT 6

AAI61115

ID AAI61115 standard; cDNA; 1727 BP.

XX AC AAI61115;

XX XX

DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 5104.

XX XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;

KW Chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW leukaemia; ss.

KW XX

XX OS Homo sapiens.

XX XX

XX PN W0200153312-A1.

XX XX

XX PD 26-JUL-2001.

XX XX

XX PF 26-DEC-2000; 2000WO-US034263.

XX XX

XX PR 23-DEC-1999; 99US-00471275.

XX PR 21-JAN-2000; 2000US-00488725.

XX PR 25-APR-2000; 2000US-00552317.

XX PR 20-JUN-2000; 2000US-00598042.

XX PR 19-JUL-2000; 2000US-00620312.

XX PR 03-AUG-2000; 2000US-00653450.

XX PR 14-SEP-2000; 2000US-00662191.

XX PR 19-OCT-2000; 2000US-00693036.

XX PR 29-NOV-2000; 2000US-00727344.

XX XX

XX PA (HYSE-) HYSEQ INC.

XX XX

XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

XX PI Zhou P, Goodrich R, Drmanac RT;

XX XX

XX DR WPI; 2001-442253/47.

XX DR P-PSDB; AAM41959.

XX XX

XX PT Novel nucleic acids and polypeptides, useful for treating disorders such

XX PT as central nervous system injuries.

XX XX

XX PS Claim 1; SEQ ID NO 5104; 10078pp; English.

XX XX

XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the

XX CC encoded polypeptides (AAM38642-AAM42213) with nootropic,

XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful

XX CC in gene therapy. A composition containing a polypeptide or polynucleotide

XX CC of the invention may be used to treat diseases of the peripheral nervous

XX CC system, such as peripheral nervous injuries, peripheral neuropathy and

XX CC localised neuropathies and central nervous system diseases, such as

XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

XX CC utilisation of the activities such as: Immune system suppression.

XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

XX CC assays for receptor activity, arthritis and inflammation, leukaemias and

XX CC C.N.S disorders. Note: The sequence data for this patent did not form

## CC part of the printed specification

XX SQ Sequence 1727 BP; 306 A; 569 C; 519 G; 332 T; 0 U; 1 Other;  
Query Match 29.2%; Score 603.8; DB 4; Length 1727;  
Best Local Similarity 98.9%; Pred. No. 4.3e-101;  
Matches 608; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1024 GGCATGTGGATCCGCTCTACCTGACCGCGCCCAAGTGTGCAAGAGACGAGACG 1083  
Db |||||

Qy 1084 TGCAACAGCCCGCTGTGGGGCAAGGACAGCCTCATCTTTGGGGCCATCACCCTGCTT 1143  
Db |||||

Qy 1144 ACGGGATTTCTGGCGTGGTTCACGGGGGAGGAGCCACGGCTGGTCCGCTGAAGACC 1203  
Db |||||

Qy 1204 CAGCGGCGCCACCACTGGTGTGCGTGGGCATGCTGGGCTCTGCCATCTTTCATCTGC 1263  
Db |||||

Qy 1264 CTGATCTTCGTGGCTGCCAAGACGACATCGTAGGAGCCTATATCTGTATCTTTCGTCGGG 1323  
Db |||||

Qy 1324 GAGACGCTGCTTTTCTAACTGGGCGCATCTGCAGACATCCTCATGTAGTGGTCAATC 1383  
Db |||||

Qy 1384 CCCACGCGCGCCCACTGCGGTGGCTTGCAGAGCTTACCTCCCACTGCTGGGGGAC 1443  
Db |||||

Qy 1444 GCCGGAGCCCTTACCTCATTTGCTTTATCTCAGACCTGATCGCCAGACACTAAGGAC 1503  
Db |||||

Qy 1504 TCCCGCTCTGGAGTTCCTGAGCCCTGGGCTACGGCTCATGCTTCGCCCTTTCGTCGTG 1563  
Db |||||

Qy 1564 GTCCTGGCGGCATGTTCTTCTCGCCACTGCGCTCTTCTTCGTGACGACCGCGCCAGG 1623  
Db |||||

Qy 1624 GCTGAGCAGCACCTG 1638  
Db |||||

Qy 615 GCTGACACGACGAGTG 629  
Db |||||

ABQ74263  
ID ABQ74263 standard; cDNA; 1587 BP.

XX AC ABQ74263;

XX DT 14-OCT-2002 (first entry)

XX DE Human 46455 transporter protein encoding cDNA SEQ ID NO:6.

XX KW Human; transporter protein; anorectic; antidiabetic; anti-Parkinsonian;

XX KW nootropic; neuroprotective; hypotensive; antidepressant; neuroleptic;

XX KW cardiovascular; immunosuppressive; gene therapy; sugar homeostasis;

XX KW obesity; diabetes; anorexia; central nervous system disorder; depression;

XX KW CNS disorder; Parkinson's disease; Alzheimer's disease; hypertension;

XX KW autonomic function disorder; schizophrenia; learning disorder; amnesia;

XX KW memory disorder; age-related disorder; cardiovascular disorder;

XX KW ischaemia reperfusion injury; restenosis; hormonal disorder;

XX KW hypothyroidism; hyperthyroidism; immune disorder;

XX KW chronic mucocutaneous candidiasis; gene; ss.



Key	Location/Qualifiers
1..1587	
/*tag= a	
/product= "46455 transporter protein"	
WO200255701-A2.	
18-JUL-2002.	
17-DEC-2001; 2001WO-US049060.	
15-DEC-2000; 2000US-0256240P.	
18-DEC-2000; 2000US-0256588P.	
21-DEC-2000; 2000US-0258028P.	
(MILL-) MILLENNIUM PHARM INC.	
Curtis RAJ;	
WPI; 2002-590672/63.	
P-FSDB; ABP52155.	
New 67076, 67102, 44181, 67084FL and 67084alt nucleic acids and proteins useful in e.g. treating disorders characterized by insufficient or excessive production of the polypeptides, e.g. cardiovascular or immune disorders.	
Claim 1; Fig 8; 364pp; English.	
The present sequence encodes the human 46455 protein, which is a member of the transporter family. The transporter family proteins (1) of the present invention have anorectic, antidiabetic, anti-Parkinsonian, nootropic, neuroprotective, hypotensive, antidepressant, neuroleptic, cardiovascular and immunosuppressive activities, and can be used in gene therapy. The human transporter proteins from the present invention are designated 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL and 67084alt. The polynucleotides encoding (1) can be used in screening assays (e.g. chromosome mapping, tissue typing, or in forensic biology), predictive medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical trials, or pharmacogenetics), as surrogate markers, and in methods of treatment (e.g. therapeutic or prophylactic). (1) are useful for treating disorders characterized by insufficient or excessive production of 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL or 67084alt substrates or production of transport 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL or 67084alt inhibitors. Disorders associated with 8099 and 46455 include disorders associated with sugar homeostasis such as obesity, diabetes or anorexia. Disorders associated with 54414, 53763, 67076, 67102, 44181, 67084FL or 67084alt are e.g. CNS disorders (e.g. Parkinson's disease or Alzheimer's disease), autonomic disorders (e.g. hypertension, depression or schizophrenia), or learning or memory disorders (e.g. amnesia or age-related disorders), cardiovascular disorders (e.g. ischaemia reperfusion injury, or restenosis), hormonal disorders (e.g. hypothyroidism or hyperthyroidism), or immune disorders (e.g. chronic mucocutaneous candidiasis)	
Sequence 1587 BP; 218 A, 529 C; 479 G; 361 T; 0 U; 0 Other;	
Query Match	27.2%; Score 563.8; DB 6; Length 1587;
Best Local Similarity	64.3%; Pred. No. 8.7e-94;
Matches 879; Conservative	0; Mismatches 482; Indels 6; Gaps 2
282	GGGCCCCCGCGCTCAGCAGCCCAACCGCGCAGCTTGGGCGCGGGGGGCGAGCCGC 341
126	GGACGAGAGGGGTGCGGCCATCACCGGCCTGTCTCCCGCGCGCTCTCATAGT 185
342	CGCCATCTCAGCTTGGGCGACGCTGCTCAACTACTCGGACAGTACACCGTGGCAGGGCT 401
186	GGCGGTGTGTGTACATCAATCTCTTGAACACTACATGACCCGCTTACCGTGGCTGGCGT 245
402	CTCTTGACATCCAGCAGCATTCTGGGGTCAAGGACCGAGGCGCGGCTCTCTGCAATC 461

Dd 1326 TGACGGCTGCGCGGAACCTGCGCCCCCTCTCTTGTGCGAGTTCCGGGCTCTGCAGTT 1385  
Qy 1536 CGCGCTCATGCTCTGCGCTTTCGTCGTGCTGCGCGGCATGTTCTTCTCGCCACTGCG 1595  
Dd 1386 CTCGCTCATGCTCTGCGCGTTTGTGGGCACTGGGCGCGCAGCTTCTTGGGACCGC 1445  
Qy 1596 GCTCTTCTGTCAGCGACCGCGGCTGAGGAGCCTGAGCAGCCTGGGG 1642  
Dd 1446 CATCTTCAATTGAGGCGCGACCGCGCGGCGGACAGCTGCACGTGCAGG 1492

RESULT 8

ADD37479  
ID ADD37479 standard; cDNA; 1587 BP.

XX AC ADD37479;  
XX DT 15-JAN-2004 (first entry)  
XX DE Human transporter 46455 cDNA #2.  
XX KW Human; ss; gene; transporter; cytostatic; anorectic; antidiabetic;  
KW anticonvulsant; gene therapy; PGC-1 associated disorder; liver tumour;  
KW obesity; epilepsy; diabetes.  
XX OS Homo sapiens.  
XX PN US2003143675-A1.  
XX PD 31-JUL-2003.  
XX PF 22-MAY-2002; 2002US-00154419.  
XX PR 12-MAY-2000; 2000US-0204211P.  
XX PR 29-JUN-2000; 2000US-0215376P.  
XX PR 31-JUL-2000; 2000US-0217859P.  
XX PR 19-SEP-2000; 2000US-0233790P.  
XX PR 25-SEP-2000; 2000US-0235107P.  
XX PR 05-OCT-2000; 2000US-0238336P.  
XX PR 14-NOV-2000; 2000US-0248364P.  
XX PR 15-NOV-2000; 2000US-0248878P.  
XX PR 15-DEC-2000; 2000US-0256240P.  
XX PR 18-DEC-2000; 2000US-0256588P.  
XX PR 21-DEC-2000; 2000US-0258028P.  
XX PR 22-JAN-2001; 2001US-0263169P.  
XX PR 14-MAY-2001; 2001US-00858194.  
XX PR 29-JUN-2001; 2001US-00895811.  
XX PR 31-JUL-2001; 2001US-00919781.  
XX PR 19-SEP-2001; 2001US-00957664.  
XX PR 25-SEP-2001; 2001US-00964295.  
XX PR 05-OCT-2001; 2001US-00972724.  
XX PR 14-NOV-2001; 2001US-00002769.  
XX PR 17-DEC-2001; 2001US-00024623.  
XX PR 22-JAN-2002; 2002US-00055025.

(MILL-) MILLENNIUM PHARM INC.

Curtis RAJ, Glucksmann MA, Meyers RE;

WPI; 2003-851783/79.

P-PSDB; ADD37478.

XX New isolated nucleic acid, useful for preparing a composition for  
PT treating PGC-1 associated disorders e.g. liver tumors, obesity, epilepsy  
PT or diabetes.

XX Claim 1; SEQ ID NO 56; 663pp; English.

XX The invention relates to an isolated nucleic acid comprising a cDNA  
CC encoding a human transporter protein, or its complement, a sequence that  
CC is 60 % identical to the cDNA, a fragment comprising at least 30  
CC nucleotides of the cDNA, or a sequence encoding a fragment of the  
CC polypeptide comprising at least 10 contiguous amino acid residues of the

CC cDNA. Also included are a vector comprising the novel nucleic acid  
CC molecule, producing the polypeptide, the isolated transporter  
CC polypeptide, an isolated antibody that specifically binds to the  
CC polypeptide, detecting the presence of the polypeptide or nucleic acid in  
CC a sample, a kit, identifying a compound that binds to, or that modulates  
CC the activity of, the polypeptide, and modulating the activity of the  
CC polypeptide. The nucleic acid is useful for preparing a composition for  
CC treating PGC-1 (not defined) associated disorders e.g. liver tumors,  
CC obesity, epilepsy or diabetes. The present sequence encodes a novel human  
CC transporter protein.

XX SQ Sequence 1587 BP; 218 A; 529 C; 479 G; 361 T; 0 U; 0 Other;

Query Match 27.2%; Score 563.8; DB 10; Length 1587;

Best Local Similarity 64.3%; Pred. No. 8.7e-94;

Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;

Qy 282 GGGCCCCGGGCTCAGCAGCCCAACCGCCAGCTTGGGCGCGGGGGGAGCGCG 341

Dd 126 GGACGAGGGGCTGCAGCGCATCACCGGCTGTCTCCCGCGCTTCGCTCATAGT 185

Qy 342 CGCATCTCTCAGCTTGGGCAACGCTCACTACCTGACAGGTACACCGTGGCAGGGT 401

Dd 186 GCGGCTGTGTCTACATCAATCTCTGAACTACATGACCGCTTACCGTGGCTGGCGT 245

Qy 402 CCTTCTGACATCCAGCAGCAGCTTTGGGGTCAAGGCGCGAGCGCGCTCTGCAGTC 461

Dd 246 CCTTCCGACATCGAGCAGTTCTTCAACATCGGGGACAGTAGCTCTGGCTCATCCAG 305

Qy 462 AGTGTTCATCTAGCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 521

Dd 306 CGTGTTCATCTCCAGTTACATGCTGTTGGCACTGTCTGTTGGCTTACCTGGTGACAGTA 365

Qy 522 CAACAGGAAGTGATCTCAGCTGCGGCATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 581

Dd 366 CAATCGGAAGTATCTCATGTGCGGGGCTATGCTTCTGCTGCTGCTGCTGCTGCTGCT 425

Qy 582 CTCCTTCATTTCCCGCAGCAGTACTTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 641

Dd 426 ATCTTTCATCCCGGAGAGCATTTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 485

Qy 642 CGGGAGGCGCAGTACTTCCACATGCGCCCACTATCATTTGGGACCTCTTTCACAAAGAA 701

Dd 486 CGGGAGGCGCAGTATTTCCACCATGCGGCCACTCTCATTTGCCGACCTCTTTGTGGCGGA 545

Qy 702 CACGCTACGCTCATGCTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 761

Dd 546 CCAGCGGAGCGGATGCTCAGCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 605

Qy 762 CTACATTTACTTGGCTCCAGCGTGAAGCAGGAGCGGAGAGCTGGCAGCTGGGCGATTGGCGGGT 821

Dd 606 CTACATTTGAGGCTCTCCAAAGTGAAGATATGGCTGGAGACTGGCAGCTGGGCTCTGAGGGT 665

Qy 822 GTCCCTGTCTCTGGGATGATCAAGGAACACTCATCTCTCATTTTGGTCCCGAGCACTAA 881

Dd 666 GACACCGGCTTAGGAGTGTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 725

Qy 882 AAGGGGTATGCGGACGAGCTCGGGA---CGAGCTCAAGCGCGGAGCTCATGGCTCCG 938

Dd 726 AAGGGGAGCGCTGGAGCGGCACCTCAGATTTTGGCACCCCTGAACCCACCTCTGCTGGTGGGC 785

Qy 939 AGATATGAAGCGCTGATTTCGAAACCGCAGCTACGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 998

Dd 786 AGATCTGAGGCTCTGGCAAGAAATCTAGTTTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 845

Qy 999 TGTCTCTCTCGCACCGGGGCGCTGGGCACTGTGGATCCCGCTCTTACCTGCACCGCGGCCA 1058

Dd 846 TGTGGCTTTGTCACTGGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 905

Qy 1059 AGTTGTGCA---GAAAGACAGCAGAGCTGCAACAGCCCGCTCTGTGGGCGCAAGGACAG 1115

Dd 906 CGTGGTCTCTGGGGAGAGACCCACCTGCTTCCCGGAGACTCTCTGCTCTTCTCTCTCTCTGACAG 965

QY 1116 CCTCATCTTTGGGCGCATCCTCTGTTACGGGATTTCTGGGCGTGTTCACGGGGCAGG 1175  
DB 966 TCTCATCTTTGGACTCATCCTGCTGACCGGAGTCTCTGGGTGTGGGCGTGGGTGGA 1025  
QY 1176 AGCCAGCGCTGTGGCGCTGAGAGCCAGCGGGCGGACCACTGGGTGTGGCGTGGG 1235  
DB 1026 GATCAGCCCGGCTCGGCACTCCACCCCGGGGTGATCCCTGGTCTGTGGCACTGG 1085  
QY 1236 CATGCTGGGCTCTGCGCATCTTCATCTGCTGATCTTCGTGGGTGCAAGAGCAGATGCT 1295  
DB 1086 CCTCTGGGCTCTGCACT 1145  
QY 1296 AGAGGCTATATCTGATCTCTGTCGGGAGAGCGTCTGTTTCTTAAGCTGGGCACTCAC 1355  
DB 1146 GGCCACTTATATTTTCATCTTCAATGGAGAGACCTCTCTCTCAAGAACTGGGCCATGCT 1205  
QY 1356 TGCAAGACATCTCATCTGCTGATCTATCCCAAGCGCGGCGGCACTGCGGTGGCTTGCA 1415  
DB 1206 GGCCGACATCTGCTGATCTGATCTGATCTCTACCCGACGCTCCACCGCGGAGGCTTCCA 1265  
QY 1416 GAGCTTCACCTCCCACTGCTGGGGAGCGCGGGAGCCCTACCTCATTTGGCTTTATCTC 1475  
DB 1266 GATCGTGTCTGCTCCACTGCTGCTGATCTGGGAGCCCTACCTCATTTGGCTGATCTC 1325  
QY 1476 AGACCTGATCCGCGAGCACTAAGCACTCCCGCTCTGGGAGTCTCTGAGCTGGGCTA 1535  
DB 1326 TGACCGGCTCGCGCGAACTGGCCCCCTCTCTTGTCCGAGTTCGGGCTCTGAGTT 1385  
QY 1536 CGGCTCATCTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1595  
DB 1386 CTGCTCATGCTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1445  
QY 1596 GCTCTTTCTGCTCAGCAGCGCGCCAGGGCTGAGCAGCACTGGGGG 1642  
DB 1446 CATCTTATTGAGCGCGACCGCGCGGCGGACAGCTGCACGTGCAGG 1492

RESULT 9

AD127958  
ID AD127958 standard; cdna; 1587 BP.

AC AD127958;

XX AD127958 (first entry)

DT 06-MAY-2004 (first entry)

XX Human 46455 cDNA coding region.

XX Human; 46455; gene; ss; ion channel family; ICF; cancer; leukaemia;  
KW Alzheimer's disease; Parkinson's disease; multiple sclerosis; epilepsy;  
KW hepatic disorder; cardiovascular disorder; cytostatic; neuroprotective;  
KW neutropic; antiparkinsonian; hepatotropic; cardiovascular.

XX Homo sapiens.

XX US2003165891-A1.

XX 04-SEP-2003.

XX 15-MAY-2002; 2002US-00146733.

XX 29-FEB-2000; 2000US-00515520.

XX 29-FEB-2000; 2000US-0185938P.

XX 03-MAR-2000; 2000US-00518866.

XX 07-APR-2000; 2000US-0195734P.

XX 11-APR-2000; 2000US-0195993P.

XX 26-APR-2000; 2000US-0199799P.

XX 19-SEP-2000; 2000US-0233537P.

XX 25-SEP-2000; 2000US-0235018P.

XX 25-SEP-2000; 2000US-0235059P.

XX 18-DEC-2000; 2000US-0256588P.

XX 21-DEC-2000; 2000US-0258028P.

XX 28-FEB-2001; 2001US-00796720.

PR 06-APR-2001; 2001US-00828035.  
PR 11-APR-2001; 2001US-00833081.  
PR 25-APR-2001; 2001US-00843128.  
PR 19-SEP-2001; 2001US-00957683.  
PR 25-SEP-2001; 2001US-00964252.  
PR 25-SEP-2001; 2001US-00964256.  
PR 17-DEC-2001; 2001US-00024623.  
XX (MILL-) MILLENNIUM PHARM INC.  
PA Curtis RAJ, Glucksmann MA, Silos-Santiago I;  
PI P-PSDB; ADI27957.  
XX WPI; 2004-069000/07.  
XX TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615, HNMDA-1, TWIK-9,  
PT alpha2delta-4, 54414, or 53763 nucleic acids and proteins, useful for  
PT preventing, diagnosing and treating e.g. cancers, Alzheimer's disease and  
PT cardiovascular disorders.  
XX Example 1; SEQ ID NO 51; 638pp; English.

PS The invention relates to TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611,  
XX IC47615, HNMDA-1, TWIK-9, alpha-2delta-4, 54414 and 53763 ion channel  
CC family (ICF) nucleic acids and proteins. The TWIK-6, TWIK-7, IC23927,  
CC TWIK-8, IC47611, IC47615, HNMDA-1, TWIK-9, alpha-2delta-4, 54414 and  
CC 53763 ICF nucleic acids and proteins may be used for preventing,  
CC diagnosing and treating ICF-related diseases. The sequences may be used  
CC to treat disorders associated with decreased expression by rectifying  
CC mutations or deletions in a patient's genome that affect the activity of  
CC ICF proteins by expressing inactive proteins or to supplement the  
CC patients own production of ICF proteins. The proteins may also be used as  
CC antigens in the production of antibodies against ICF proteins and in  
CC assays to identify modulators of ICF protein expression and activity. The  
CC anti-ICF protein antibodies, agonists and antagonists may be used to  
CC regulate ICF protein expression and activity. The antibodies may also be  
CC used as diagnostic agents for detecting the presence of ICF proteins in  
CC samples (e.g. by immunosay). The nucleic acids and proteins may be used  
CC to prevent, diagnose and treat a wide variety of disorders, e.g. cancers  
CC and leukaemia, Alzheimer's disease, Parkinson's disease, multiple  
CC sclerosis, epilepsy, hepatic disorders and cardiovascular disorders. This  
CC sequence represents cDNA encoding the human 46455 protein of the  
CC invention. Note: The sequence data for this patent is also available in  
CC electronic format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 1587 BP; 218 A; 529 C; 479 G; 361 T; 0 U; 0 Other;

SQ Query Match 27.2%; Score 563.8; DB 12; Length 1587;  
Best Local Similarity 64.3%; Pred. No. 8.7e-94;  
Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;

QY 282 GGGCCCCGGCTCAGCAGCCCAACCGGCCAGCTGGGCCGGGGGGGCGAGCCGC 341  
DB 126 GGACACAGAGGGGCTGCAGCGCATACCGGCTGTCTCCGGCCGCTTCGGCTCATAGT 185  
QY 342 CGCCATCTCTCAGCTTGGGCAACGCTGTCACTACCTGGACAGGTACACCGTGACGCGCT 401  
DB 186 GGGGGTGTGTGTATCATCAATCTCTGAATACATGACCGCTTACCGTGGCTGGCGT 245  
QY 402 CTTTCTGGACATCCAGCAGCATTTGGGGTCAAGACCCAGAGCGCGGCTGTGTCAGTTC 461  
DB 246 CTTTCCGACATCGAGCAGTCTTCAACATCGGGGACAGTAGCTCTGGGCTCATCCAGAC 305  
QY 462 AGTGTTTCATCTGTAGCTTTCATGTGGCTGCCCCCATCTTCGGGTACTCTGGCGACCGCTT 521  
DB 306 CGTGTTTCATCTCCAGTTATAGTGTGTGGACCTGTGTGGTACTCTGGTGACAGGTA 365  
QY 522 CAACAGAGAGGTGATTCTCAGCTGCGGCATTTTCTTCTGGTGGCGCGTCACTTCTCCAG 581  
DB 366 CAATCGGAAGTATCTCATGTGCGGGGCACTTGGCTTCTGTCTCCCTGTGACACTGGGGTC 425  
QY 582 CTCCTTCATTCGCCAGCAGTACTTCTGGCTGTGGTCTGTCCCGGGGGCTGTGGGTCAT 641

Db 426 ATCTTATCCCGAGAGCATTTCTGGTCTCTCTGACCCGGGCTGTGGGGT 485  
Qy 642 CGGGAGCCAGCTACTCCACATCGCCCCCACTATCATTTGGCGACTTTCACCAAGAA 701  
Db 486 CGGGAGCCAGTATTATCCACATCGCCCACTCTCATTTGCCGACCTTTTGTGGCGA 545  
Qy 702 CACGGTACGTCTGCTCGCTTCTTACTTGGCCATCCCACTGGGCACTGGCTGGG 761  
Db 546 CCAGCGGACCGGATGCTCAGCATCTTCTATTTGGCCATTCGGTGGGCACTGGTGGG 605  
Qy 762 CTACATTACTGGCTCCAGCGTGAGCAGCGAGCGAGCTGGCACTGGGCACTGGGCGT 821  
Db 606 CTACATTGACGGCTCCAAAGTGAGGATATGGCTGGAGACTGGCACTGGGCTCTGAGGGT 665  
Qy 822 GTCCCCCTGTCTGGGCACTGATCACAGGAACACTCATCTCATTTCTGGTCCAGCCACTAA 881  
Db 666 GACACCGGCTCTAGAGTGGTGGCGTTCTGCTGCTTCTGTTCTGTTAGTGGGAGCGCC 725  
Qy 882 AAGGGTCACTCCGACCACTGGGGA---CGAGTCAAGCCCGGACCTCATGGTCCG 938  
Db 726 AAGGGGAGCCGTGGAGCGCCACTCAGATTTGCCACCCCTGAACCCCACTCTGTTGGGC 785  
Qy 939 AGATATGAAGCCCTGATTTCAACCGCAGCTAGCTCTTCTCTCCCTGGCCACGTGGC 998  
Db 786 AGATCTGAGGCTCTGGCAAGAAATCTAGTTTCTGCTCTGTTCTTCCCTGGGCTTCACTGC 845  
Qy 999 TGTCTCTTCCCACTGGGCGCTGGGCACTGGATCCGCTCTACCTGACCCGCGCCA 1058  
Db 846 TGTGGCTTTGTACAGGCTCTCTGGCTCTGTTGGCTCCGCACTTCTGCTGGTCCCG 905  
Qy 1059 AGTTGTGCA---GAAGACAGCAGAGAGTGCAACAGCCCGCCCTGTGGGGCCAAAGGACAG 1115  
Db 906 CGTGGTCTTGGGAGAGACCCACCTGCTTCCCGGAGACTCTCTCTCTCTGACAG 965  
Qy 1116 CCTCATCTTGGGCGCACTACTGCTTTACGGGATTTCTGGGCTGTGTCAGGGGCGAG 1175  
Db 966 TCTCATCTTTGATCTCATCTGCTGACCGGAGTCTGGGCTGTGGGCTGTGGG 1025  
Qy 1176 AGCCACGCGTGTGGCGCTGAGAGACCCAGCGGCGCCACCTGTTGTGTGGCGTGGG 1235  
Db 1026 GATCAGCGCGGCTCGGCACTCCACCCCGGGCTGATCCCTGTTGTGTGGCACTGG 1085  
Qy 1236 CATCTGGGCTCTGCCATTTTCATCTGCTGATTTGCTGGCTGCCAAGACAGCATCGT 1295  
Db 1086 CCTCTGGGCTCTGACCTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1145  
Qy 1296 AGGAGCTTATCTGTATCTTCTGCGGGAGACGCTGCTGTTTCTTAACCTGGGCACTAC 1355  
Db 1146 GGCCACTTATATTTTATCTTCTATTTGGAGAGACCTCTCTGCTTCAATGAGCTGGGCACTCGT 1205  
Qy 1356 TGACAGACATCTCATGTACGTGGTCTATCCCAAGCGGCGCCACTGCGTGGCTTGA 1415  
Db 1206 GGCCGACATTTCTGTACGTGGTGTATCCCTACCGAGCTTCCACCGCGGAGGCTTCCA 1265  
Qy 1416 GAGCTTCACTTCCACCTGCTGGGGGAGCGCGGAGCCCTTACCTCATTTGGCTTTTATCTC 1475  
Db 1266 GATGCTGTGCTCCCACTGCTGGGTGATGCTGGAGCCCTTACCTCATTTGGCTGATCTC 1325  
Qy 1476 AGACTGATCCGCGAGAGCACTAAGGACTCCCGCTCTGGAGTTCTGAGCCTGGGCTA 1535  
Db 1326 TGACCGCTGCGCGGAACTGGCCCCCTCTTCTTGTTCGAGTTCTGGGCTCTGCACTG 1385  
Qy 1536 CGGCTCATGCTGCGCTTCTGCTGCTGCTGCTGCGGAGATGTTCTTCTGCGCACTGC 1595  
Db 1386 CTCCTCATGCTCTGCGGTTTGTGGGCACTGGGCGGCGAGCCCTTCTTGGGCACTGC 1445  
Qy 1596 GCTCTTCTGTCAGCGACCGCGGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1642  
Db 1446 CATCTTCAATTGAGCGCGAGCCCGCGGCGGACAGCTGCAGCTGCAGG 1492

RESULT 10  
ADL61203

ID ADL61203 standard; DNA; 2165 BP.  
XX AC  
XX ADL61203;  
DT 03-JUN-2004 (first entry)  
XX Human protein tyrosine kinase biomarker spinster-like DNA.  
DE predictor set; protein tyrosine kinase; cytostatic; antiangiogenic;  
KW vasotrophic; vulnery; pharmacogenomic; drug sensitivity; breast cancer;  
KW hypervascular disease; angiogenesis; wound healing scar; human;  
XX biomarker; ds; gene; spinster-like.  
OS Homo sapiens.  
XX WO2004020583-A2.  
XX 11-MAR-2004.  
XX 26-AUG-2003; 2003WO-US026491.  
XX 27-AUG-2002; 2002US-0406385P.  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX Huang F, Han X, Reeves KA, Amler L, Fairchild CR, Lee FY;  
PI Shaw P;  
XX WPI; 2004-239171/22.  
DR P-PSDB; ADL61327.  
XX New predictor sets with a plurality of polynucleotides and/or  
PT polypeptides whose expression pattern predicts cell response to a  
PT compound that modulates protein tyrosine kinase activity, useful in  
PT treating breast cancer.  
XX Claim 2; SEQ ID NO 127; 649pp; English.  
XX The invention relates to a novel predictor set comprising a plurality of  
CC polynucleotides and/or polypeptides whose expression pattern is  
CC predictive of the response of cells to treatment with a compound that  
CC modulates protein tyrosine kinase activity or members of the protein  
CC tyrosine kinase pathway. The molecules of the invention demonstrate  
CC cytostatic, antiangiogenic, vasotrophic and vulnery activities and may  
CC be useful in the field of pharmacogenomics, in particular for determining  
CC drug sensitivity and in treating breast cancer, hypervascular diseases,  
CC angiogenesis and scars in wound healing. The current sequence is that of  
CC a human protein tyrosine kinase biomarker DNA of the invention.  
XX SQ Sequence 2165 BP; 308 A; 729 C; 662 G; 466 T; 0 U; 0 Other;  
Query Match 27.28; Score 563.8; DB 12; Length 2165;  
Best Local Similarity 64.3; Pred. No. 9e-94;  
Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;  
Qy 282 GGGCCCCGGGCTCAGCAGCCCAACCGGCGAGCTTGGCGCGGGGGGCGAGCGC 341  
Db 450 GGACGAGGCGGTGAGCGCATCCCGGCTGTCTCCCGCGCTTCCGCTCTCATAGT 509  
Qy 342 CGCCATCTCAGCTTGGGCAACGCTGCTCAACTACCTGAGCAGGTACACCGTGGCGGCT 401  
Db 510 GGCGGTGCTGTGCTACATCAATCTCTGAACTACATGACCGCTTACCGTGGCTGGGT 569  
Qy 402 CCTTCTGGACATCCAGCAGCACTTGGGTCAAGGACCGAGGCGCGGCTGTGCGATC 461  
Db 570 CCTTCCGACATCGAGGAGTTCTTCAACATCGGAGCAGTAGTCTCGGCTCATCCAGAC 629  
Qy 462 AGTGTTCATCTGTAGCTTTCATGTGGCTGCCCATCTTGGCTTACCTGGGCGACCGCTT 521  
Db 630 CGTGTTCATCTCAGTTACATGTTGGGACCTGTGTTGGCTTACCTGGTGTGACAGTA 689  
Qy 522 CAACAGGAGGTGATTTCTCAGCTGCGGCAATTTTCTTGTGTCGGCGCTACCTTCTCCAG 581

Db 690 CAATCGGAAGTATCTCATGTGCGGGGCATTTGCCCTTCTGGTCCCTGGTGACACTGGGGTC 749  
Qy 582 CTCCTTTCATTTCCCGACGACTTCTGGGTCTGGTCCCTGGTCCCGGGGCTGGTGGGCAT 641  
Db 750 ATCTTTCATCCCGGAGAGATTTCTTGGCTCTCTCTCTGACCCCGGGGCTGGTGGGGT 809  
Qy 642 CGGGAGGCGAGCTACTTCCACCATCCGCCCCCACTATCATTTGGCGACTCTTTCACCAAGAA 701  
Db 810 CGGGAGGCGAGTTATTCACCATCCGGCCCACTCTCATTTGGCGACTCTTTGTGGCCGA 869  
Qy 702 CAGCGTAGCGTCACTGTCTGTCTTCTTACTTTCGCTCCCACTCCCACTGGGAGTGGCTGGG 761  
Db 870 CCAGCGGAGCGGATGCTCAGCATCTTCTACTTTGCAATTCGGTGGGAGTGGTCTGGG 929  
Qy 762 CTACATTTAGTCTCAGCGTGAAGCAGGCGAGCGGAGACTGGCACTGGGCAATGGGGT 821  
Db 930 CTACATTTGAGGCTCAAGAGTGAAGGATATGGCTGGAGACTGGCACTGGGCTCTGAGGT 989  
Qy 822 GTCCCTCTCTGGGATGATCACAGGAACACTCATCTCTTCTGGTCCCACTCAAA 881  
Db 990 GACACCGGTCTAGGAGTGGTGGCGCTTCTGTCTGTCTTCTGGTAGTGGGAGCGCC 1049  
Qy 882 AAGGGTCTATGCCGACAGCTGGGGA---CCAGCTCAAGGCCCGGACCTCATTTGGTCCG 938  
Db 1050 AAGGGAGCGCTGGAGCGCACTCAGATTTGGCCACCCCTGAACCCCACTCGTGGTGGGC 1109  
Qy 939 AGATATGAAGGCCCTGATTCGAACCGCAGCTACGTCTTCTCTCTCGCTGGCCACGCTCGG 998  
Db 1110 AGATCTGAGGGCTCTGGCAAGAAATCCTAGTTTGGCTCTTCTCTCTGGGCTTCACTGC 1169  
Qy 999 TGTCTCTCTTCCGCAACCGGGGCGCTGGGCACTTGGATCCCGCTCTACTGTGACCGCCCA 1058  
Db 1170 TGTGGCTTTTGTACCGGCTCTCTGTCTGTGGGCTCGGCACTTCTGTGGTTCGCG 1229  
Qy 1059 AGTTGTGCA---GAAGACACAGAGAGAGCTGCAACAGCCCGCTGTGGGGCCCAAGACAG 1115  
Db 1230 CGTGGTCTTGGGGAGACCCACCTCTCCCGGAGACTCTCTCTTCTCTGACAG 1289  
Qy 1116 CTTATCTTTTGGGGCATCATCTCTTACGGGATTTCTGGGGCTGGTCAAGGGGCGAG 1175  
Db 1290 TCTCATCTTTTGGACTCATCACCTGCTGACCGGAGTCTTGGGTGGGGCTGGGTGGA 1349  
Qy 1176 AGCCAGCGCTGTGGCTGAGGCTGAAGACCCAGCGGGCGGACCCCACTGGTGTGGCTGGG 1235  
Db 1350 GATCAGCGCGGCTCGGCCACTCCAAACCCCGGCTGATCCCTGTGTGTGCCACTGG 1409  
Qy 1236 CATGTGGGCTTGGCATCTTCAATCTGCTGTATCTTGTGGGTGCCAAGAGAGCATGCT 1295  
Db 1410 CTTCTGGGCTTGCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1469  
Qy 1296 AGGAGCTATATCTGTATCTTCTGCGGGAGAGCGTCTGTTTCTTAACTGGGCCATCAG 1355  
Db 1470 GGCCATTTATTTTATCTTCTTATGGAGAGACCTTCTGTTCATGAATGGGCCATGCT 1529  
Qy 1356 TGCAGACATCTCATGTAGTGTGATCCCAAGCGGGCGGCGGCGGCTGGCTTGA 1415  
Db 1530 GGCGGACATCTGCTGTAGTGTGATCTCTACCCAGCGCTCCAGCGCGGAGGCTTCCA 1589  
Qy 1416 GAGCTTTCACCTCCACCTGCTGGGGACGCGGGAGCCCTTACCTCATTTGGCTTTATCTC 1475  
Db 1590 GATCGTGTCTGCCACCTGCTGGGTGATGTGGGAGCCCTTACCTCATTTGGCTGATCTC 1649  
Qy 1476 AGACCTGATCCGCGAGACACTAAGGACTCCCGCTCTGGGAGTCTTCTGAGCTGGGCTA 1535  
Db 1650 TGACCCGCTCGCGCGAACTGGCCCCCTCTTCTTGTTCGAGTTCCTGGGCTCTGAGTT 1709  
Qy 1536 CGGGCTCATGCTCTGGCTTCTGCTGTGTCTTGGGCGGATGTTCTTCTCTGCCACTGCG 1595  
Db 1710 CTGCTCATGCTCTGGCGTTTGTGGGCACTTGGGCGGCGGCGGCTTCTTGGGCAACCG 1769  
Qy 1596 GCTCTTTCTTGTGAGGACCGCGCCAGGGCTGAGCAGCACTTGGGG 1642  
Db 1770 CATCTTTCATTTAGGCGGACCGCGCGGGGCGACAGCTGCAGTGCAGG 1816

RESULT 11

AAI57853  
ID AAI57853 standard; cDNA; 2218 BP.

XX

AC AAI57853;

XX 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 56.

DE Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.

XX Homo sapiens.

OS

XX WO200153312-A1.

PN

XX 26-JUL-2001.

XX

XX 26-DEC-2000; 2000WO-US034263.

XX

XX 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-00552317.

PR 20-JUN-2000; 2000US-00598042.

PR 19-JUL-2000; 2000US-00620312.

PR 03-AUG-2000; 2000US-00653450.

PR 14-SEP-2000; 2000US-00662191.

PR 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

PA

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Zhou J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

PI Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX P-PSDB; AAM38697.

DR

XX Novel nucleic acids and polypeptides, useful for treating disorders such

PT as central nervous system injuries.

XX

PS Claim 1; SEQ ID NO 56; 10078pp; English.

XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the

CC encoded polypeptides (AAM38642-AAM42213) with nontropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders. Note: The sequence data for this patent did not form

CC part of the printed specification

XX

SQ Sequence 2218 BP; 318 A; 745 C; 683 G; 472 T; 0 U; 0 Other;

XX

Query Match 27.2%; Score 563.8; DB 4; Length 2218;

Best Local Similarity 64.3%; Pred. No. 9e-94;

Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;



CC of the transporter family. The transporter family proteins (1) of the  
CC present invention have anorectic, antidiabetic, anti-Parkinsonian,  
CC noreptropic, neuroprotective, hypotensive, antidepressant, neuroleptic,  
CC cardiovascular and immunosuppressive activities, and can be used in gene  
CC therapy. The human transporter proteins from the present invention are  
CC designated 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL and  
CC 67084alt. The polynucleotides encoding (1) can be used in screening  
CC assays (e.g. chromosome mapping, tissue typing, or in forensic biology),  
CC predictive medicine (e.g. diagnostic assays, prognostic assays,  
CC monitoring clinical trials, or pharmacogenetics), as surrogate markers,  
CC and in methods of treatment (e.g. therapeutic or prophylactic). (1) are  
CC useful for treating disorders characterised by insufficient or excessive  
CC production of 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL or  
CC 67084alt substrates or production of transport 8099, 46455, 54414, 53763,  
CC 67076, 67102, 44181, 67084FL or 67084alt inhibitors. Disorders associated  
CC with 8099 and 46455 include disorders associated with sugar homeostasis  
CC such as obesity, diabetes or anorexia. Disorders associated with 54414,  
CC 53763, 67076, 67102, 44181, 67084FL or 67084alt are e.g. CNS disorders  
CC (e.g. Parkinson's disease or Alzheimer's disease), autonomic function  
CC disorders (e.g. hypertension, depression or schizophrenia), or learning  
CC or memory disorders (e.g. amnesia or age-related disorders),  
CC cardiovascular disorders (e.g. ischaemia reperfusion injury or  
CC restenosis), hormonal disorders (e.g. hypothyroidism or hyperthyroidism),  
CC or immune disorders (e.g. chronic mucocutaneous candidiasis)  
XX  
SQ Sequence 2230 BP; 318 A; 752 C; 688 G; 472 T; 0 U; 0 Other;

Query Match 27.2%; Score 563.8; DB 6; Length 2230;  
Best Local Similarity 64.3%; Pred. No. 9e-94;  
Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;

QY 282 GGGCCCCGGCTCAGACGCCAAGCCGCGCAGCTGGGCGCGGGGGGCGAGCCG 341  
DB 501 GGACCGAGAGGGCTCGACGCGATCACCGGCGCTGTCTCCCGCGGTTCATAGT 560

QY 342 CGCATCTCTCAGCTTGGGCAACGCTGCTCAACTACCTGGACAGGTACACCGTGGCAGCGCT 401  
DB 561 GGGGGTGTGTCTACATCAATCTCTGAACTACATGGACCGCTTCACTGGTGGCGT 620

QY 402 CTTTCTGGACATCCAGACGACATTTGGGGTCAAGGACCGAGGCGCGCGCTGTGAGTC 461  
DB 621 CTTTCCGACATCGAGCAGTCTTCAACATCGGGGACAGTAGCTCTGGGCTATCCAGAC 680

QY 462 AGTGTTCATCTGAGCTTCAATGCTGCTGCCCATCTTCCGCTACCTGGGACCGCTT 521  
DB 681 CGTGTTCATCTCAGTTACATGCTGTGGCACCCTGTGTGGTACTCTGGTGACAGTA 740

QY 522 CAACAGGAAGGTGATTTCTCAGCTGCGGCAATTTCTTCTGCTCGGCGGTCACTTCTCCAG 581  
DB 741 CAATCGGAAGTATCTCATGTGCGGGGCAATTCCTTCTGCTCCTGTGACACTGGGGTC 800

QY 582 CTCCTTCATTTCCGAGAGTACTTCTGGCTGTGCTCTGCTCCCGGGGCTGTGGGCAT 641  
DB 801 ATCTTTTCATCCCGGAGAGATTTCTGGCTGTCTCTCTGACCCCGGGGCTGTGGGGT 860

QY 642 CGGGAGGGGAGCTACTTCCACCATCGCCCGCCTATCATTTGGGAGCTCTTTCACCAAGAA 701  
DB 861 CGGGAGGGGAGCTATTTCCACCATCGGGCCCATCTCATTTGCCGACCTCTTTGTGGCGA 920

QY 702 CAGCGGTACGCTCATGCTGTCCGTCTTCTACTTTGGCCATCCCACTGGGAGTGGCTGGG 761  
DB 921 CCAGCGGAGCGGATGCTCAGCATCTTCTACTTTGGCATTCCTGGTGGGAGTGTCTGGG 980

QY 762 CTACATTATCGGCTCGAGCTGAAGCGGAGCGGAGACTGGCAGTGGGCATTTGGGGT 821  
DB 981 CTACATTGCAAGGCTCCAAAGTGAAGGATATGGCTGGAGACTGGCAGTGGGCTCTGAGGT 1040

QY 822 GTCCCTGTCTCGGCGATGATACAGGAACATCATCTCATTTCTGCTGCCAGCCACTAA 881  
DB 1041 GACACGGGCTTAGAGTGTGGCGCTTCTGCTGTCTTCTGTGTGTGGTGGGAGCGGCC 1100

QY 882 AAGGGGTCAATGCCGACAGCTCGGGGA---CCAGCTCAAGGCCGAGCTCATGTGCTCCG 938  
|||||

DB 1101 AAGGGAGCGCTGGAGCGCCACTCAGATTTTGCCACCCCTGAACCCCACTCTGTGTGGGC 1160  
QY 939 AGATATGAAGCCCTGATTCGAAACCGCAGCTACGCTTCTCTCTCCCTGGCCACGCTCGGC 998  
DB 1161 AGATCTGAGGCTCTGGCAAGAAATCTAGTTCCTCTCTCTCTCTCTCTCTCTCTCTCT 1220  
QY 999 TGTCTCTCTTGGCCACGCGGGGCTCTGGGATCTGGATCCCGCTCTACTCTGACCGGCCA 1058  
DB 1221 TGTGGCTTTGTACGGGCTCCCTGGCTCTGGGCTCGGATCTCTCTCTCTCTCTCTCT 1280  
QY 1059 AGTTGTGCA---GAAGACAGCAGAGACGTGCMAACAGCGCGCTCTGTGGGCGCAAGACAG 1115  
DB 1281 CGTGTCTCTTGGGAGAGACCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1340  
QY 1116 CTTATCTTTGGGGGCATCACTGCTTTACGGGATTTCTGGGCGTGGTCAACGGGGGCGAG 1175  
DB 1341 TCTATCTTTGGACTCATACCTGCTGACCGGAGTCTTGGGTGTGGGCTGGGTGTGGA 1400  
QY 1176 AGCCACGCGCTGTGTCGCGCTTGAAGACCCAGCGGCGCGACCTGGTGTGTGCCGTGGG 1235  
DB 1401 GATCAGCGCGGCTCCGCCACTCCACCCCGGGCTGATCCCTGGTCTGTGCCACTGG 1460  
QY 1236 CATGCTGGGCTCTGCACTTCTCATCTGCTGATCTTCTGTGGGCTGCCAAGACAGCATCGT 1295  
DB 1461 CTTCTGTGGGCTCTGCACCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1520

QY 1296 AGGACCTATATCTGATCTTCTGTCGGGAGACGCTGCTGTTTCTTAACCTGGGCCATCAC 1355  
DB 1521 GGCCACTTATATTTTCTCTCTTCTTGGAGAGACCTCTCTGTCATGAACCTGGGCCATCGT 1580

QY 1356 TGCAGACATCTCTATGATGCTGATCCCAACGCGGCGCGCACCTGGCTGGCTTTGCA 1415  
DB 1581 GGCCGACATCTGCTGTAGCTGTGATCCCTACCGAGGCTCCACCGCGAGGCTTCCA 1640

QY 1416 GAGCTTACCTCCACTGCTGGGGAACGCGGAGGCGGCTACCTCATTTGGCTTTATCTC 1475  
DB 1641 GATCGTGTGTCCTACCTGCTGGGTGATGCTGGGAGCCCTACCTCATTTGGCTGATCTC 1700

QY 1476 AGACCTGATCCGCGAGACACTAAGACCTCCCGCTCTGGGAGTTCCTGAGCCTGGGCTA 1535  
DB 1701 TGACCGCTTGGCGCGGAACCTGGCCCTCTCTCTTGTCCGAGTTCGCGGCTTGAGTT 1760

QY 1536 CGCGCTCATGCTCTGCGCTTTCTGCTGTGCTCTGGGCGGCAATGTTCTTCTCTGCCACTGC 1595  
DB 1761 CTGCTCATGCTCTGCGGCTTTTGGGCACTTGGGCGCGAGCCCTTCTCTGGGCACTGC 1820

QY 1596 GCTCTTCTGTCAGGACCGCGGCGGCTGAGCAGCAGCTGGGG 1642  
DB 1821 CATCTTCAATTGAGGCGGACCGCGGCGGCGAGCTGCACGTCAGG 1867

RESULT 13  
ADD37477 standard; cDNA; 2230 BP.  
XX  
AC ADD37477;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Human transporter 46455 cDNA #1.  
XX  
KW Human; ss; gene; transporter; cytostatic; anorectic; antidiabetic;  
KW anticonvulsant; gene therapy; FGC-1 associated disorder; liver tumour;  
KW obesity; epilepsy; diabetes.  
XX  
OS Homo sapiens.  
XX  
PN US2003143675-A1.  
XX  
PD 31-JUL-2003.  
XX  
PF 22-MAY-2002; 2002US-00154419.  
XX





1821 CATCTTCATTGAGCGCGACCGCGCGGCGACAGCTGCAGTGCAGG 1867

RESULT 14

AD127956  
ADI27956 standard; cDNA; 2230 BP.

AD127956;

06-MAY-2004 (first entry)

Human 46455 cDNA.

Human; 46455; gene; ss; ion channel family; ICF; cancer; leukaemia;  
Alzheimer's disease; Parkinson's disease; multiple sclerosis; epilepsy;  
hepatic disorder; cardiovascular disorder; cytostatic; neuroprotective;  
neotopic; antiparkinsonian; hepatotropic; cardiovascular.

Homo sapiens.

US2003165891-A1.

04-SEP-2003.

15-MAY-2002; 2002US-00146733.

29-FEB-2000; 2000US-00515520.

29-FEB-2000; 2000US-0185938P.

03-MAR-2000; 2000US-00518866.

07-APR-2000; 2000US-0195734P.

11-APR-2000; 2000US-0195933P.

26-APR-2000; 2000US-0199799P.

19-SEP-2000; 2000US-0233537P.

25-SEP-2000; 2000US-0235018P.

15-DEC-2000; 2000US-0256240P.

18-DEC-2000; 2000US-0256588P.

21-DEC-2000; 2000US-0258028P.

28-FEB-2001; 2001US-00796720.

06-APR-2001; 2001US-00828035.

11-APR-2001; 2001US-00833081.

25-APR-2001; 2001US-00843128.

19-SEP-2001; 2001US-00957683.

25-SEP-2001; 2001US-00964252.

17-DEC-2001; 2001US-00964256.

17-DEC-2001; 2001US-00024623.

(MILL-) MILLENNIUM PHARM INC.

Curtis RAJ, Gluckmann MA, Silos-Santiago I;

WPI; 2004-069000/07.

P-PSDB; ADI27957.

TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, HNMDA-1, TWIK-9,

alpha2delta-4, 54414, or 53763 nucleic acids and proteins, useful for

preventing, diagnosing and treating e.g. cancers, Alzheimer's disease and

assays to identify modulators of ICF protein expression and activity. The  
anti-ICF protein antibodies, agonists and antagonists may be used to  
regulate ICF protein expression and activity. The antibodies may also be  
used as diagnostic agents for detecting the presence of ICF proteins in  
samples (e.g. by immunoassay). The nucleic acids and proteins may be used  
to prevent, diagnose and treat a wide variety of disorders, e.g. cancers  
and leukaemia, Alzheimer's disease, Parkinson's disease, multiple  
sclerosis, epilepsy, hepatic disorders and cardiovascular disorders. This  
sequence represents cDNA encoding the human 46455 protein of the  
invention. Note: The sequence data for this patent is also available in  
electronic format from USPTO at seqdata.uspto.gov/sequence.html.

Sequence 2230 BP; 318 A; 752 C; 688 G; 472 T; 0 U; 0 Other;

Query Match 27.2%; Score 563.8; DB 12; Length 2230;

Best Local Similarity 64.3%; Pred. No. 9e-94; Mismatches 482; Indels 6; Gaps 2;

Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;

282 GGGCCCGCGCTCAGCAGCCCAACCGCGCAGCTTGGGCCCGGGGGGCGACGCCG 341

501 GGACGAGGAGGGCTGAGGGCATCACCGGCTGTCTCCGGCGTTCGGCTCTCATGT 560

342 OGCCATCTCAGCTTGGGCAAGCTGCTCAATCTCTGACAGAGTACACCGTGGCGGCT 401

561 GCGGCTGTGTGTACATCAATCTCTGAACCTACATGACCGCTTACCGTGGTGGCGT 620

402 CCTCTGGACATCCAGCAGCACTTGGGGTCAAGGACGAGGCGCGGCTCTCTGAGTC 461

621 CCTTCCGACATCGAGCAGTCTTCAACATCGGAGCAGTAGTCTGGGCTCATCCAGAC 680

462 AGTGTTCATCTGTAGCTTTCATGGTGGCTGCCCCCATCTTTCGGCTACCTGGGGCGACGCTT 521

681 CGTGTTCATCTCCAGTTACATGGTGTGGCACCTGTGTGTGGCTACCTGGGTGACAGTGA 740

522 CAACAGGAGTGATTTCTCAGTGGCGCATTTTCTTGTGTGGCGGCTCACCTTCTCCAG 581

741 CAATCGAAGTATCTCATGTGGGGGCAATTCCTTCTGTCTCTGTGTGACATGGGGTCTC 800

582 CTCCTTCATTTCCAGCAGTACTTCTGGTGTGTGTCTCTCCGGGGGCTGGTGGGCT 641

801 ATCTCTTCATCTCCCGGAGAGCATTTCTGGCTGTCTCTCTGACCCGGGGGCTGGTGGGCT 860

642 CGGGAGGCGCAGTACTCTCACCATCGCCCCCATCTATCTATTTGGTCCCGGAGTAA 701

861 CGGGAGGCGCAGTATTTCCACCATCGGCCCATCTCTATTTGGCGGCTCTTTTGGCGCA 920

702 CACGGTACGCTCATGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 761

921 CCAGCGGAGCGGATGCTCAGCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 980

762 CTACATTTAGGCTCCAGGCTGAAGCGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCT 821

981 CTACATTTAGGCTCCAAAGTGAAGATATGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCT 1040

822 GTCCCTCTCTCTGGGCTGATCAGAGAACCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 881

1041 GACACGGGCTTAGAGTGGTGGCGCTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1100

882 AAGGGTCTATGCGGACCGAGCTCGGGGA---CCAGCTCAAGGCCCGGACCTCATGGCTCCG 938

1101 AAGGGGAGCGGTGAGCGCCACTCAGATTTGACACCCCTGAAACCCACCTCGTGTGTGGG 1160

939 AGATATGAAGGCTTATTCGAACCGGAGCTACGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 998

1161 AGATCTGAGGCTCTGGCAAGAAATCTTAGTTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1220

999 TGCTCTCTCTCTGGGCGGCTTGGGATCTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1058

1221 TGTCGCTTGTGACGGGCTCTCTGTCTGTGGCTCGGCTCTCTCTCTCTCTCTCTCTCTCT 1280

1059 AGTTGTGCA---GAGACAGCAGAGAGCTGCAACAGCGCGCTGTGGGGCCAGAGACAG 1115

1281 CGTGTCTCTTGGGGAGACCCCAACCTTCCCTCCCGAGAGACTCTCTCTCTCTCTCTCTCT 1340

QY 1116 CCTCATCTTTGGGCGCAATACCTGCTTTTACGGGATTTCTGGGCGTGTGTCACGGGGCAGG 1175  
DB TCTCATCTTTGGGCGCAATACCTGCTTTTACGGGATTTCTGGGCGTGTGTCACGGGGCAGG 1400  
QY 1176 AGCCAGCGGCTGGGCGGCTGAAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1235  
DB GATCAGCGGCGGCTGGGCGGCTGAAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1460  
QY 1236 CATGCTGGGCTGCTGCGCTTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1295  
DB CTCTGGGCTGCTGCGCTTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1520  
QY 1296 AGGAGCTTATCTGTATCTTCTGCTGGGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1355  
DB GGCACCTTATATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1580  
QY 1356 TGCAGACATCTCATGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1415  
DB GGCAGACATCTCATGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1640  
QY 1416 GAGCTTCACTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1475  
DB GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1700  
QY 1476 AGACCTGATCGCGCAGACCACTAAGGAGCTCCCGCTCTGCGAGTTCCTGAGCCTGGGCTA 1535  
DB TGACCGGCTGCGCGGAGCACTGCGCGCTCTCTTCTGCTGCGAGTTCGCGGCTGCTGAGTT 1760  
QY 1536 CGCGCTCATGCTGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1595  
DB CTCGCTCATGCTGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1820  
QY 1596 GCTCTTCTGCTGAGCGG 1642  
DB CATCTTCTGAGCGG 1867

## RESULT 15

ID ABL90674/c  
XX ABL90674 standard; cDNA; 2166 BP.  
AC ABL90674;  
XX  
DT 24-MAY-2002 (first entry)  
XX Human polynucleotide SEQ ID NO 1236.  
DE  
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;  
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200190304-A2.  
XX  
PD 29-NOV-2001.  
XX  
PF 18-MAY-2001; 2001WO-US016450.  
XX  
XX 19-MAY-2000; 2000US-0205515P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Birse CE, Rosen CA;  
PI  
XX WPI: 2002-122018/16.  
DR  
DR P-PSDB; ABB90265.  
XX  
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
PT prevention of neural, immune system, muscular, reproductive,

gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders.  
XX  
PS Claim 4; SEQ ID NO 1236; 2081pp + Sequence Listing; English.  
XX  
CC The invention relates to novel genes (ABL9449-ABL90853) and proteins (ABB9040-ABB9044) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pat\_sequences

XX Sequence 2166 BP; 464 A; 646 C; 715 G; 336 T; 0 U; 5 Other;

Query Match 27.2%; Score 562.6; DB 6; Length 2166;  
Best Local Similarity 64.1%; Pred. No. 1.5e-93;  
Matches 876; Conservative 3; Mismatches 482; Indels 6; Gaps 2;  
QY 282 GGGCCCCGGCGCTCAGCAGCCCAACCGGCCAGCTGGGGCGGGGGGCGAGCGCGC 341  
DB 1754 GRACAGAGGGGGCTGCGAGCGCATCMCGGCCCTGTCTCCCGCGCTCTCATAGT 1695  
QY 342 CGCCATCTCAGCTTGGGCAACGTGCTCACTACCTGGAGGATACCTGGAGCGGT 401  
DB 1694 GCGCGTGTGTGTATCAATCTCTGAATCATATGAGACCGCTTACCGTGGTGGCGT 1635  
QY 402 CCTTCTGGACATCCAGCAGCACTTTGGGGTCAAGGACCGAGCGCGCGCTGTGCAATC 461  
DB 1634 CCTTCCGACATCGAGCAGTCTTCAACATCGGGGACAGTAGTCTGGGCTCATCCAG 1575  
QY 462 AGTGTTCATCTGAGCTTTCATGCTGGTGTGCGGCTGCGGCTACCTGGGCGGCGGCTT 521  
DB 1574 CGTGTTCATCTCCAGTTACATGCTGTGGCACCTGTGTGGTGTGCTGCTGCTGCTG 1515  
QY 522 CAACAGAGGTGATCTCAGCTGCGGCACTTTCTTGTGTGGCGGTACCTTCTCCAG 581  
DB 1514 CAATCGGAAGTATCTCATGTGCGGGGCGCATTTGCTTCTGTGCTGCTGCTGCTGCT 1455  
QY 582 CTCCTTCATTTCCAGCAGTACTTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 641  
DB 1454 ATCTTCTATCCCCGAGAGCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1395  
QY 642 CGGGAGGGCGAGCTACTTCCACCATCGCCCCCACTATCATTTGGGCGGCGCTTTCACCA 701  
DB 1394 CGGGAGGGCGAGTTATTCACCATCGGCCACTCTCTCATTTGCGGCGGCTTCTTGG 1335  
QY 702 CACCGTAGCTCATGCTGCTGCTTCTTCTTCTGCGGATCCCACTGGGCGAGTGGCGTGG 761  
DB 1334 CCAGCGGAGCGGATGCTCAGCATCTTCTACTTTGCGATTTCCGCTGGGCGAGTGGT 1275  
QY 762 CTACATTACTGGCTCCAGCGTGAAGCAGCGGCGGAGCTGGACCTGGGCGATTGGCGGT 821  
DB 1274 CTACATTGAGGCTCCAAAGTGAAGATATGGCTGGAGACTGGGCGCTCTGAGGGT 1215  
QY 822 GTCCCTGTCTGGGCGATGATCACAGGAACATCATCTCTCATTTGGTCCGAGCACTAA 881  
DB 1214 GACACCGGCTTAGGAGTGTGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1155  
QY 882 AGGGGTCTATGCCGACAGCTCGGGA---CGAGTCAAGGCCCGGAGCTCATGCTCGC 938  
DB 1154 AAGGGGAGCGGTGGAGCGGCACTCAGATTTGGCCACCCCTGAAACCCACCTCGTGGTGG 1095







EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 26.2%; Score 542.8; DB 4; Length 2201;  
Best Local Similarity 64.3%; Pred. No. 7.3e-101;  
Matches 877; Conservative 1; Mismatches 478; Indels 8; Gaps 4;

```
QY 285 CCCCGGGCTCAGACGCCAACCGCGAGCTTGGCGCGGGGGGGGCGAGCGCGGC 344
D 416 CCAGAGGGCTGCGAGCGCATCAMCGGCGCTCTCCCGGCGGTTCGGCTCTCATAGTGGC 475
QY 345 CATCCTCAGCTTGGGCGCAAGTGCTCAACTACTGGACAGGTACACCGTGGCGAGCGTCC 404
D 476 GGTGCTGTACATCAATCTCTGAATACATGACCGCTTACCGTGGCTGGCGTCT 535
QY 405 TCTGGACATCAGACGACATTTGGGGTCAAGGACGAGGCGCGGCTGTCTGACAGTCA 464
D 536 TCCGACATCGACGAGTCTTCAACATCGGGGACAGTAGTCTCTGGGCTCATCAGACCGT 595
QY 465 GTTCATCTGAGTCTCATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 524
D 596 GTTCATCTCAGTTCATAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 655
QY 525 CAGGAAGTGATCTCAGCTGCGGCATTTCTCTGTGCGGCGTCACTTCTCCAGCTC 584
D 656 TCGGAAGTATCTCATGTGCGGGGCAATGCGCTCTGTGCTGCTGCTGCTGCTGCTGCTG 715
QY 585 CTTCAATCCCGAGAGTACTTCTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 644
D 716 CTTCAATCCCGAGAGATTTCTGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 775
QY 645 GGAGGCGAGTACTCAGACATCGGCCCGCACTATCATTTGGGACCTCTTACCAAGAACAC 704
D 776 GGAGGCGAGTATTTCCACCACATCGGCCCGCACTCTCATTTGGGACCTCTTGTGGCGGCA 835
QY 705 GCGTACGCTCATGTGTCGCTTCTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 764
D 836 GCGGAC -CGATGCTCAGATCTTCTACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
QY 765 CATTAATGCTCCAGCGTGAAGCAGGAGCGGAGACTGGCACTGGGCACTTGGGGTGTG 824
D 895 CATTTGACAGGCTCCAAAGTGAAGATATGGCTGGAGACTGGCACTGGGCTCTGAGGGTGAC 954
QY 825 CCCTGCTCTGGGCAATGATCAGAGAACATCATCTCTCATTTGCTGCTCCAGCCACTAAAG 884
D 955 ACCGGGTCTAGGAGTGTGGCGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
QY 885 GGGTCATGCCAGCAGCTCGGGA --CGAGCTCAGGCGCGGACCTCATGGCTCCGAGA 941
D 1015 GGGAGCGTGGAGCGCCACTCAGATTTGGCACCCCTGAACCCCACTCTGGTGGGCGAGA 1074
QY 942 TATGAAGCGCTGATTTGAAACCGCAGCTACGCTCTCTCTCTCTGCGCCACGCTCGGCTGT 1001
D 1075 TCTGAGGCTCTGGCAGAAATCTTAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1134
QY 1002 CTCCTTCCGACGCGGGGCGCTGGGCAATGAGATCCCGCTCTACTCTGACCGCGCCCAAGT 1061
D 1135 GGGCTTTGTACGCGGCTCCCTGGCTCTGTGGGCTCCGGCATTCCTGTGGTTCCTGCTGCTG 1194
QY 1062 TGTGCA --GAAGACGACGAGCGTGCACAGCCCGCTGTGGGCGCAAGACAGCT 1118
D 1195 GGTCTTGGGAGACCCCAACCTGCTCCCGAGACTCTGCTGCTCTCTCTGACAGTCT 1254
QY 1119 CATCTTTGGGGCCATCACCTGCTTTACGGATTTCTGGGCTGTGTACGCGGCGAGGAGC 1178
D 1255 CATCTTTGGACTCATCACTTGCCTGACCGAGTCTCTGGGCTGTGGGCTGTGGGCTGTGAGAT 1314
QY 1179 CAGCGCTGTGCGGCTGTAAGACCCAGCGGGCGGACCCACTGCTGTGTGCTGCGGCGAT 1238
```

```
D 1315 CAGCCGCGGCTCCGCCACTCCAAACCCCGGGCTGATCCCTGCTGTGTGCACTGGGCT 1374
QY 1239 GCTGGGCTCGCCATCTTTCATCTGCTGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1298
D 1375 CTTGGGCTCTGCACTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1434
QY 1299 AGCCTATATCTGATCTTCTGCGGGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1358
D 1435 CACTTATATTTTATCTTTCATTTGGAGAGACCTCTCTCTCTGCTGCTGCTGCTGCTGCTG 1494
QY 1359 AGACATCTCATGTACGTGCTATCCCAACCGCGGGCGGCACTGCTGCTGCTGCTGCTGCTG 1418
D 1495 CGACATCTCTGTACGTGCTGATCCCTACCCGACGCTCCACCGCGGCGCTTCCAGAT 1554
QY 1419 CTTCACTCCACCTGCTGCGGGAGCGCGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1478
D 1555 CGTGTCTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1614
QY 1479 CTTGATCCGCGAGACACTAAGACTCCCGCTCTGCGGAGTCTGAGCCTGGGCTACGC 1538
D 1615 CCGCTGCGCGGAACTGGCGCCCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1674
QY 1539 GCTCATGCTCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1598
D 1675 GCTCATGCTCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1733
QY 1599 CTTCTTCTGTCAGCGACCGCGCGGCTGAGCAGCAGCTGCGGG 1642
D 1734 CTTCTTCTGAGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1777
```

## RESULT 2

US-09-149-476-246  
Sequence 246, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
CURRENT FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,617  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23



[illegible]

D <sub>b</sub>	780	GAGATCAGCGCGGYTCCGCCACTTCAACCCCGGGCTGATCCCTGGTCTGTGCCACT	839
Q <sub>y</sub>	1234	GGCATGTGGGCTGCACATTTTCATCTGCCTGATCTTCTGGCTGCAGAGCAGCATC	1293
D <sub>b</sub>	840	GGCTCTCGGGCTGTGACCCCTTCTTCTCTGTCCCTTGCCCTGGCCCGTGGTAGCATC	899
Q <sub>y</sub>	1294	GTAGGACCTATATCTGTATCTTCGTCGGGGAGCGTGTCTTTTCTAACTGGGCCATC	1353
D <sub>b</sub>	900	GTGGCCACTTATATTTTTCATCTTCATTTGGAGAGACCCTCTGTCCATGAACTGGGCCATC	959
Q <sub>y</sub>	1354	ACTGCAGACATCTCATGTAGTGGTCATCCCCACGGCGGGCGGCCACTGCCGTGGCCCTTG	1413
D <sub>b</sub>	960	GTGGCCGACATTTGTGTGACGTGGTGATCCCTACCGAGGCTCCACCGCGGAGGCCTTC	1019

1414	CAGAGCTTCACCTCCACACCTGCTGGGGAGACGCCGGGAGACCCCTACCTTCATTTGGCTTTATC	1417
1020	CAGATCGTGTGCTCCCAACCTGCCTGGGTGATGCTGGGAGCCCTACCTCATTTGGCTGATC	1079
1474	TCAGACCTGATCCGCGACAGACACTAAGGACTCCCCCGCTCTGGGAGATTCTTGAGCCTGGGC	1533
1080	TCAGACCGCTCGCGCGGAACTGGCCCCCTCTCTTTGTGCCGATTCGGGGCTTCGAG	1139
1534	TAGCGCTCATGTCTGCCCCCTTTTCGTGCTGGTCTCGGGCGGCATGTTCTTTCTCTCGCCACT	1593
1140	TTTCGCTCATGCTCTGCGGTTTGTGGGSCACTGGGCGGCGCACT-TTCCTGGGCACC	1198
1594	CGCGTCTTCTTCTGTCAGCGACCGGGCCAGGGCTGAGCAGCACCTCGGGG	1642
1199	GNCATCTTCATTAGGCGCGACCGCGCGGSCACAGCTGCACGTGCAGG	1247

RESULT 3  
 US-09-149-476-87  
 ; Sequence 87, Application US/09149476  
 ; Patent No. 6420526  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: 186 Human Secreted proteins  
 ; FILE REFERENCE: P2002P1  
 ; CURRENT APPLICATION NUMBER: US/09/149,476  
 ; CURRENT FILING DATE: 1998-09-08  
 ; EARLIER APPLICATION NUMBER: PCT/US98/04493  
 ; EARLIER FILING DATE: 1998-03-06  
 ; EARLIER APPLICATION NUMBER: 60/040,162  
 ; EARLIER FILING DATE: 1997-03-07  
 ; EARLIER APPLICATION NUMBER: 60/040,333  
 ; EARLIER FILING DATE: 1997-03-07  
 ; EARLIER APPLICATION NUMBER: 60/038,621  
 ; EARLIER FILING DATE: 1997-03-07  
 ; EARLIER APPLICATION NUMBER: 60/040,626  
 ; EARLIER FILING DATE: 1997-03-07  
 ; EARLIER APPLICATION NUMBER: 60/040,334  
 ; EARLIER FILING DATE: 1997-03-07  
 ; EARLIER APPLICATION NUMBER: 60/040,336  
 ; EARLIER FILING DATE: 1997-03-07  
 ; EARLIER APPLICATION NUMBER: 60/040,163  
 ; EARLIER FILING DATE: 1997-03-07  
 ; EARLIER APPLICATION NUMBER: 60/047,600  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,615  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,597  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,502  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,633  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,583  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,617  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,618  
 ; EARLIER FILING DATE: 1997-05-23

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89</											

EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 21.9%; Score 453; DB 4; Length 1460;

Best Local Similarity 65.2%; Pred. No. 9,4e-83;

Matches 716; Conservative 0; Mismatches 375; Indels 8; Gaps 3;

QY 550 ATTTCCTTCGTCGGCGCTCACCTTCCTCCAGCTCCTTCATTCCTCCAGCAGTACTTCTGG 609  
DB 1 ATTGCTTCCTGCTCCCTGCTGACACTGGGGTCACTCTTCATCCCGGAGAGCAATTCCTGG 60  
QY 610 CTGCTGTCCTGTCCTCGGGGCTGTCGGGCATCGGGAGGCCAGCTACTCCACCATGCC 669  
DB 61 CTGCTCCTCTGACCCGGGGCTGTCGGGGTTCGGGGAGGCCAGTATTATCCACCATGCCG 120  
QY 670 CCCACTATCATTCGGCAGCTCTTCACCAAGAACACGGTACGCTCATGCTCTCGCTCTTC 729  
DB 121 CCCACTCATTCGGCAGCTCTTCGGCCAGCAGCGGACCGGATGCTCAGCATCTTC 180  
QY 730 TACTTCGCCATCCCACTGGGCGAGTGGCTGGGCTACATTAATGCTCCAGCGTCAAGCAG 789  
DB 181 TACTTCGCAATTCGGTGGGCGAGTGTCTGGGCTACATTCGAGCTCCAAAGTGAAGAT 240  
QY 790 GCACCCGAGACTGGGCACTGGGCATTCGGGGTGTCCCTGTCTCTGGGAGATCAGAGGA 849  
DB 241 ATGCTGGAGACTGGGCACTGGGCTCTGAGGGTGACACCGGGTCTAGGAGTGTGGCGGTT 300  
QY 850 ACATCATCTCATTCCTGCTCCAGCCACTAAAGGGGTATGCGGACCGAGCTCGGGA- 908  
DB 301 CTGCTGCTGTCCTGCTAGTGGGAGCGCCAGGGAGCGCTGGAGCGGCACTCAGAT 360  
QY 909 --CCAGCTCAAGGCCCGGACCTCATGCTCCAGATATGAAGGCCCTGATTCGAAACCGC 966  
DB 361 TTGCCACCCCTGAACCCACCTGCTGTGTGGGCGAGATCTGAGGGCTCTGGCAAGAACTCT 420  
QY 967 AGCTAGCTCTTCCTCCCTGGCCACGCTGGCTGTCTCTTCGCGACCGGGGCTCTGGGC 1026  
DB 421 AGTTTCGCTCTCTTCCTGGGCTTCACCTGCTGTGGGCTTGTGTCAGGGCTCTCCCTGGCT 480  
QY 1027 ATGTGGATCCGCTCTACCTGCACCGCGCCCAAGTTGTC---AGAGACAGCAGAGACG 1083  
DB 481 CTGTGGGCTCCGGCAATTCCTGCTGCTGCTCCGCTGCTGCTGGGAGACCCCAACCTGC 540  
QY 1084 TGCAACAGCCCGCTGTCGGGGCCCAAGCAGACCTCATCTTTGGGGCCATCAGCTGCTTT 1143  
DB 541 CTTCCCGGAGACTCTGCTCTTCTCTGACAGTCTCATCTTTGAGCATCATCAGCTGCTG 600  
QY 1144 ACGGATTTCTGGCGTGTGTCACGGGGGCGAGGAGCCACGGCTGTGGTCCGCCCTGAAGACC 1203  
DB 601 ACCGAGTCTCTGGTGTGGGCTGCTGGGTGTGGAGATCAGCGCGGCTCCGCCACTCCAAC 660  
QY 1204 CAGGGGCCACCACTGCTGTGTGCTGGGCACTGCTGGGCTGCGCATCTTCATCTGC 1263  
DB 661 CCGCGGCTGATCCCTGGTGTGTGCGCATCTGGGCTCTGGGCTCTGACCCCTTCCTCTTC 720  
QY 1264 CTGATCTTCGTGGCTGCAAGAGCAGCATCTGAGAGCCTATATCTGTATCTCTGTCGGG 1323  
DB 721 CTGTCCCTTCGCTGGCGGCTGTGAGCATCTGTGGCAGCTTATATTTCACTTCATTTGA 780  
QY 1324 GAGACGCTGCTGTTTCTAACTGGGCGCATCACTGCAGACATCTCTCATGTAGTGGTCATC 1383

DB 781 GAGACCTCTCTGTCCATGAACCTGGGCCCATCTGGGCCGACATTCCTGTGTACGTGTGATC 840  
QY 1384 CCCACGGCGCGCGCACTGCGGTGGCTTGCGAGAGCTTCACCTCCACCTGCTGGGGGAC 1443  
DB 841 CCTACCGAGCGCTCCACCGCGAGGCTTTCAGATGCTGTGCTCCACCTGCTGGGTGAT 900  
QY 1444 GCGGGAGCGCCCTACCTCATTTGCTTTATCTCAGACCTGATCCGCCAGAGCACTAAGGAC 1503  
DB 901 GCTGGAGCGCCCTACCTCATTTGGCTGATCTCTCAGCCGCTGCGCGGAACTGGGCCCCC 960  
QY 1504 TCCCGGCTCTGGGAGTTCCTGAGGCTGGGCTACGGGCTCAGCTGCTGCGCTTCTGCTG 1563  
DB 961 TCCTTCTGTCGAGTTCGCGGCTCTGCACTCTCTGCTCATGCTCTGCGCGTCTGCTGGG 1020  
QY 1564 GTCCTGGCGCGCATGTTCTTCTCGCCACTGCGCTCTTCTTCGTCAGGACCGCGCCAGG 1623  
DB 1021 GCATGGGGGCG--GCATTCCTCTGGGCAACCGGCACTTCATTGAGGCCGACCGCGGCG 1078  
QY 1624 GCTGAGCAGCAGCTGGGGG 1642  
DB 1079 GCACAGCTGCAGGTGACG 1097

## RESULT 4

US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZ9pt-Fls  
; US-08-232-463-14

Query Match 3.4%; Score 69.4; DB 1; Length 7218;

	Best Local Similarity	4.4%; Pred. No. 4.6e-05;	Mismatches	Indels	Gaps
	Matches	Conservative	189;	55;	0;
Qy	1621	AGGGCTGACGACACTTGGGGGAGACGCGCGGGGTCAAGGTGGTCATCAGCGGGG	1680		
	:	:::	:	:	:
	:	:::	:	:	:
	:	:::	:	:	:
Db	1337	RR	1278		
Qy	1681	CGGGCCCCGGCACTGCTTGGCACATCTGTCTGGGGGCCAGCTCACCGGAGTGCTG	1740		
	:	:::	:	:	:
	:	:::	:	:	:
	:	:::	:	:	:
Db	1277	RR	1218		
Qy	1741	GCAGGGCACTCTGTCAGCCCGGGGAGATGGAGAGCCACGGGTGGGGAGAGAGAG	1800		
	:	:::	:	:	:
	:	:::	:	:	:
	:	:::	:	:	:
Db	1217	RR	1158		
Qy	1801	AGAGAGAGTAAGAGGAAAGAGAAAGAAGTCAGAAAAGTAAGAGGAAGGGAGGGCCC	1860		
	:	:::	:	:	:
	:	:::	:	:	:
	:	:::	:	:	:
Db	1157	RR	1098		
Qy	1861	CAGCTTTGAAACCACCTAAGTCCAGAGACAACCCAAAGTCTGGATCCACAGACACC	1917		
	:	:::	:	:	:
	:	:::	:	:	:
	:	:::	:	:	:
Db	1097	RR	1041		
	:	:::	:	:	:
	:	:::	:	:	:
	:	:::	:	:	:

RESULT 5  
 US-09-616-289-48/c  
 ; Sequence 48, Application US/09616289  
 ; Patent No. 6632923  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lees, Ann M.  
 ; APPLICANT: Lees, Robert S.  
 ; APPLICANT: Law, Simon W.  
 ; APPLICANT: Arjona, Anibal A.  
 ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
 ; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
 ; TITLE OF INVENTION: ATHEROSCLEROSIS  
 ; FILE REFERENCE: 10797-004001  
 ; CURRENT APPLICATION NUMBER: US/09/616.289  
 ; CURRENT FILING DATE: 2000-07-14  
 ; PRIOR APPLICATION NUMBER: US 09/517,849  
 ; PRIOR FILING DATE: 2000-03-02  
 ; PRIOR APPLICATION NUMBER: US 08/979,608  
 ; PRIOR FILING DATE: 1997-11-26  
 ; PRIOR APPLICATION NUMBER: US 60/031,930  
 ; PRIOR FILING DATE: 1996-11-27  
 ; PRIOR APPLICATION NUMBER: US 60/048,547  
 ; PRIOR FILING DATE: 1997-06-03  
 ; NUMBER OF SEQ ID NOS: 53  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 48  
 ; LENGTH: 2561  
 ; TYPE: DNA  
 ; ORGANISM: Oryctolagus cuniculus  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (246)... (1895)  
 US-09-616-289-48

```

QY      204  GCTGTGCGGACGCTAAGCGCGGCCCCGACCGGACCCCCCGCGCACCCCCCGGCAC 257
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      587  GCGCGCGCGGCGCGCGCGGCTCGCGCGCGCGGCGCGCTGACGCGCGCGCGC 534

RESULT 6
US-09-128-155-16/c
; Sequence 16, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 152331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(152331)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-16

```

RESULT 7  
US-09-252-991A-15609  
; Sequence 15609, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 15609

; LENGTH: 1560

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-15609

Query Match 2.9%; Score 60.2; DB 4; Length 1560;  
Best Local Similarity 47.9%; Pred. No. 0.0024;  
Matches 241; Conservative 0; Mismatches 253; Indels 9; Gaps 2;

Qy 260 CCAGCTGCGAGCTACTGCAAGGGCCCCCGCGCTCAGCAGCCCAAAACCGGCCAGCTTGG 319  
Db 3 CCAGCTGCGTTCGGTTCGGTCCCGCCAGGTGCGGTCTTGTCCGATCTCGGCCAGAAC 196  
Qy 320 GCCCGGGGGGGGAGCCGCCCATCTCTGGACATCCAGCAGCACTTTGGGCTCAAGGACC 439  
Db 63 CGCCCGCGCTGCGTACCGTCTCGCGGCTGATGCTCGATCTTCTCAGCGCCCTCG 122  
Qy 380 ACAGGTACACGTGGCAGGGTCTCTTGGACATCCAGCAGCACTTTGGGCTCAAGGACC 439  
Db 123 ACCAGACCATGTCGCGGTCTCCCTGCGGCGATCTCGGCGACTTCGCGACCTCGACC 182  
Qy 440 GAGCGCGCGCTGCTCAGTCACTGTTATCTAGCTTCATGTTGCGGCTGCGCCCATCT 499  
Db 183 TGTGCGCTG---GGTGATCTCGGCTACATGTTGCGGATGACCGTATCGATGCCGATCT 239  
Qy 500 TCAGCTACCTGGGCGACCGTTCACAGAGGATGATCTCAGCTGCGGCACTTTCTTCT 559  
Db 240 ACAGCAAGCTCGGCGACTCTACGGCGCGCGCTGTGATGCTCTTCGCCATGTCGTGT 299  
Qy 560 GGTGCGCGCTACCTTCCAGCTCTTCAATCCAGCAGTACTTCTGGCTGCTGCTCC 619  
Db 300 TCACCGCGCTGCTGCTGCTTTGCGGCTGCG-----GCAGAGCATGGGCCAATGCTGC 353  
Qy 620 TGTCCCGGGGCTGTTGGGATCGGAGGCGCAGCTACTCCACATCGCCCGCCACTATCA 679  
Db 354 TGGCGGAGTGTCTCAGGGGATCGGCGGCGGCTGATGGCGGTGAGCCAGCGCATCA 413  
Qy 680 TTGGCGACCTTTCACCAAGAACACGCGTACGCTGCTCGTCTTCTTACTTCGCCA 739  
Db 414 TCGCGACATCGTCCCGCGGAGCGCGGCGGCTGATGGCGGTGAGCCAGCGCATGT 473  
Qy 740 TCCACTGGGCGAGTGGCTGGC 762  
Db 474 ACGCCATCGCAGGTGCGCGC 496

## RESULT 8

US-09-252-991A-15645  
; Sequence 15645, Application US/09252991A  
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 15645

; LENGTH: 1680

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-15645

Query Match 2.9%; Score 60.2; DB 4; Length 1680;  
Best Local Similarity 47.9%; Pred. No. 0.0024;

Matches 241; Conservative 0; Mismatches 253; Indels 9; Gaps 2;

Qy 260 CCAGCTGCGAGCTACTGCAAGGGCCCCCGCGCTCAGCAGCCCAAAACCGGCCAGCTTGG 319  
Db 137 CCAGCTGCTTCCGGTTCGGTCCCGCCAGGTGCGGTCTTGTCCGATCTCGGCCAGAAC 196  
Qy 320 GCCCGGGGGGGGAGCCGCCCATCTCTGGACATCCAGCAGCACTTTGGGCTCAAGGACC 439  
Db 197 CGCCCGCGCTGCGTACCGTCTCTCGCGGCGCTGATGCTGTGATCTTCTCAGCGCCCTCG 256  
Qy 380 ACAGGTACACGTGGCAGGGTCTCTTGGACATCCAGCAGCACTTTGGGCTCAAGGACC 439  
Db 257 ACAGACCATGTCGCGGTCTCCCTGCGGCGATCTCGGCGACTTCGCGACCTCGACC 316  
Qy 440 GAGCGCGCGCTGCTCAGTCACTGTTATCTAGCTTCATGTTGCGGATGACCGTATCGATGCGATCT 499  
Db 317 TGTGCGCTG---GGTGATCTCGGCTACATGTTGCGGATGACCGTATCGATGCGATCT 373  
Qy 500 TCAGCTACCTGGGCGACCGTTCACAGGAGGATGATCTCAGCTGCGGCACTTTCTTCT 559  
Db 374 ACAGCAAGCTCGGCGACTCTACGGCGCGCGCTGTGATGCTCTTCGCCATCTCGCTGT 433  
Qy 560 GGTGCGCGCTACCTTCCAGCTCTTCAATCCAGCAGTACTTCTGGCTGCTGCTCC 619  
Db 434 TCACCGCGCTGCTGCTGCTTTGCGGCTGCG-----GCAGAGCATGGGCCAATGCTGC 487  
Qy 620 TGTCCCGGGGCTGTTGGGATCGGAGGCGCAGCTACTCCACATCGCCCGCCACTATCA 679  
Db 488 TGGCGGAGTGTCTCAGGGGATCGGCGGCGGCTGATGGCGGTGAGCCAGCGCATCA 547  
Qy 680 TTGGCGACCTTTCACCAAGAACACGCGTACGCTGCTGCTGCTTCTTACTTCGCCA 739  
Db 548 TCGCGACATCGTCCCGCGGAGCGCGGCTGATGGCGGTGATGGCGGTGATGGCGCATGT 607  
Qy 740 TCCACTGGGCGAGTGGCTGGC 762  
Db 608 ACGCCATCGCAGGTGCGCGC 630

## RESULT 9

US-09-252-991A-15709/c

; Sequence 15709, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 15709

; LENGTH: 1899

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-15709

Query Match 2.9%; Score 60.2; DB 4; Length 1899;

Best Local Similarity 47.9%; Pred. No. 0.0025;

Matches 241; Conservative 0; Mismatches 253; Indels 9; Gaps 2;

Qy 260 CCAGCTGCGAGCTACTGCAAGGGCCCCCGCGGCTCAGCAGCCCAAAACCGGCCAGCTTGG 319  
Db 1877 CCAGCTGCTTCCGGTTCGGTCCCGCCAGGTGCGGTCTTGTCCGATCTCGGCCAGAAC 1818  
Qy 320 GCCCGGGGGGGGAGCCGCCCATCTCTAGCTTTGGGCAACGCTGCTCAACTACTCTGG 379  
Db 1817 CGCCCGCGCTGCGTACCGTCTCGCGGCTGATGCTGCTGATCTTCTCAGCGCCCTCG 1758  
Qy 380 ACAGGTACACCGTGGGCGAGCGTCTTCTGACATCCAGCAGCACTTTGGGCTCAAGGACC 439





Qy 231 GACCGGACCCCGCGGACCCCG 253  
Db 107 GGNNCTCTCCAGCGCGCGCTCG 85

## RESULT 13

US-09-616-289-45/c  
; Sequence 45, Application US/09616289  
; Patent No. 6632923  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; FILE REFERENCE: 10797-004001  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45  
; LENGTH: 1614  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)... (1614)  
US-09-616-289-45

Query Match 2.9%; Score 59.8; DB 4; Length 1614;

Best Local Similarity 52.2%; Pred. No. 0.0029; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CGCGCGCCCGATCCGCGCGCGCATGATGCTGCGAATGCGCTCGGCTCGCGCGCGCGCGG 65  
Db 540 CGCGCGCGCCAGGGCGCGCCCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 481  
Qy 66 CGCGGAGGAGGAGCGGACGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 125  
Db 480 GGGGGGCGCTGTGCG 421  
Qy 126 GGCTGGCGGTAGCGGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 185  
Db 420 GACGGGCGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 361  
Qy 186 GGGCGATGAGTGCAGACGCTGTGCGGCGAGCGTAAGCGGCGCGCGCGCGCGCGCGCG 245  
Db 360 GCGGGGCG 301  
Qy 246 CACCCCGCGCACCCC 260  
Db 300 CTGAGCGCGCGCGC 286

## RESULT 14

US-09-616-289-50/c  
; Sequence 50, Application US/09616289  
; Patent No. 6632923  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; FILE REFERENCE: 10797-004001  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 50  
; LENGTH: 12425  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-616-289-50

Query Match 2.9%; Score 59.8; DB 4; Length 12425;

Best Local Similarity 52.2%; Pred. No. 0.0045; Mismatches 133; Conservative 0; Indels 0; Gaps 0;

Qy 6 CGCGCGCCCGATCCGCGCGCGCATGATGCTGCGAATGCGCTCGGCTCGCGCGCGCGCGG 65  
Db 3371 CGCGCGCGCCAGGGCGCGCCCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 3312  
Qy 66 CGCGGAGGAGGAGCGGACGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 125  
Db 3311 GGGGGGCGCTGTGCG 3252  
Qy 126 GGCTGGCGGTAGCGGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 185  
Db 3251 GACGGGCGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3192  
Qy 186 GGGCGATGAGTGCAGACGCTGTGCGGCGAGCGTAAGCGGCGCGCGCGCGCGCGCGCG 245  
Db 3191 GGGGGGCG 3132  
Qy 246 CACCCCGCGCACCCC 260  
Db 3131 CTGAGCGCGCGCGC 3117

## RESULT 15

US-09-976-594-1091  
; Sequence 1091, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael  
; APPLICANT: Buchsinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 1091  
; LENGTH: 2030  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673549 221042.1  
US-09-976-594-1091

Query Match 2.9%; Score 59.2; DB 4; Length 2030;  
Best Local Similarity 49.1%; Pred. No. 0.004;







[illegible]

```
QY 1586 TCGCCACTGCGCTCTTCTTCGTACGACGACCGCGCCAGGGCTGAGCAGACCTG 1638
Db 1501 TCGCCACTGCGCTCTTCTTCGTACGACGACCGCGCCAGGGCTGAGCAGAGTG 1553

RESULT 3
US-10-024-623-6
; Sequence 6, Application US/10024623
; Publication No. US20020187524A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: 8099, 46455, 54414, 53763, 67076, 67102, 44181,
; TITLE OF INVENTION: 67084FL, AND 67084 ALT, HUMAN PROTEINS AND METHODS OF
; FILE REFERENCE: USE THEREOF
; CURRENT APPLICATION NUMBER: US/10/024,623
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,240
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/256,588
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/258,028
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1587
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1587)
US-10-024-623-6

Query Match 27.2%; Score 563.8; DB 13; Length 1587;
Best Local Similarity 64.3%; Pred.No. 5.6e-133;
Matches 879; Conservative 0; Mismatch 482; Indels 6; Gaps 2;

QY 282 GGGCCCCGCGCTCAGCAGCCCAAAACCGCCAGCTTGGCCCGGGGGGGGCGACCGC 341
Db 126 GGACGAGGAGGGCTCGACGCGATCACCGCGCTGTCTCCGCGGTTCGGCTCTCATAGT 185

QY 342 CGCCATCCTCAGTTGGGCAAGTGTCTCACTACCTGGACAGTACACCGTGGCGGCGT 401
Db 186 GCGGGTGTGTCTACATCAATCTCTGAATACATGAGACCGCTTACCCTGGTGGCGT 245

QY 402 CTTCTGGACATCCAGCAGCACTTTGGGGTCAAGGACCGAGCGCGCGCTGTGCGAGTC 461
Db 246 CTTTCCGACATCGACAGTCTTCAACATCGGGGACAGTAGTCTTGGGCTCATCCAGAC 305

QY 462 AGTGTTTCATCTGATGCTTCATGTTGGCTGCGCCCATCTTTCGGTACCTGGGCGACCGTT 521
Db 306 CGTGTTTCATCTCAGTTACATGTTGGCAGCTGTGTTGGCTTACCTGGGTGACAGGTA 365

QY 522 CAACAGAGGTGATTTCTCAGTGGCGGCACTTTCTCTGTTGGCGCGCTCACCTTCTCCAG 581
Db 366 CAATCGGAAGTATCTCATGTGCGGGGGCATTTGCTTCTGTTCCCTGGTACACTGGGGTC 425

QY 582 CTCTTTTCATTTCCAGCAGTACTTCTGGTGTGTTCTGTTCTCCCGGGGGCTGTGGGCAT 641
Db 426 ATCTTTCATTTCCCGAGAGCATTTCTTGGTGTCTCTCTCTCTGACCCGGGGCTGTGGGGT 485

QY 642 CGGGGAGGCGAGTACTCAGCATTCGCGCCCATCATCTATCTGTTGGGACTCTTTCACCAAGAA 701
Db 486 CGGGGAGGCGAGTATTTACCACTTCGCGCCCATCTCATTTGGCGGACTCTTTTGTGCCGA 545

QY 702 CAGCGTACGCTCATGCTGTCGCTCTTCTACTTTCGCGCATCCCATCGGAGTGGCTGG 761
Db 546 CCAGCGGAGCGGATGCTCAGCATCTTCTACTTTTGCAATTCGGTGGGAGTGGTCTGG 605

QY 762 CTACATTACTGGCTCCAGCGTGAAGCAGGCGCGGAGACTGGCACTGGGCACTGGGGT 821
Db 1587
```

```
Db 606 CTACATTGAGGCTCCAAAGTGAAGGATATGGCTGGAGACTGGCACTGGGCTCTGAGGGT 665
QY 822 GTCCCTCTGTCTGGGCAATGATCAAGNAACACTCATCTCACTTCTGGTCCAGCCACTAA 881
Db 666 GACACCGGGTCTAGGAGTGGTGGCCGTTCTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGT 725
QY 882 AAGGGGTTCATGCCGACAGCTCGGGGA---CCAGCTCAAGGCCCGGAGCCTCATGTGCTCG 938
Db 726 AAGGGAGCCGTGAGGCGCACTCAGATTTCGACCCCTGAACCCCACTCTGTGTGTGGC 785
QY 939 AGATATGAAGGCCCTGATTTCGAAACCCGAGCTACGTCTTCTCTCTCTCTCTCTCTCTCTCT 998
Db 786 AGATCTGAGGGCTCTGGCAAGAAATCTAGTTTCTGTCTCTCTCTCTCTCTCTCTCTCTCT 845
QY 999 TGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1058
Db 846 TGTGGCCCTTTGTCAAGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 905
QY 1059 AGTTGTGCA---GAAGACAGCAGAGACGTGCAACAGACCCCGCCCTGTGGGGCCCAAGACAG 1115
Db 906 CGTGTCTCTGGGAGACCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 965
QY 1116 CCTCATCTTTGGGGCCATCACCTGTTTACCGGATTTCTGGGCGTGTGTACCGGGGCGAG 1175
Db 966 TCTCATCTTTGGAATCATCATCTGCTGACCGGAGTCTCTGGGTGTGGGCTGTGGGTGTGGA 1025
QY 1176 AGCCACGCGCTGTGGCGCTGAAAGACCCAGCGGCGGCGACCTGCTGTGTGCGGTGG 1235
Db 1026 GATCAGCGCGCGGCTCGCCCACTCAACCCCGGGCTGATCCCTCTGTCTGTGCTCTGTGCT 1085
QY 1236 CATGTGGGCTCTGCCATCTTTCATCTGCTGATCTTCTGTGGTGTGCAAGAGCAGCATCGT 1295
Db 1086 CTCTCTGGGCTCTGCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1145
QY 1296 AGGAGCTATATCTGTATCTTCTGTCGGGAGACGCTGCTGTGTCTTCTAATCTGGGCTATC 1355
Db 1146 GGGCACTTATATTTTTCATCTTTCATTTGGAGAGACCTCTCTCTCATGNACTGGGCTATCGT 1205
QY 1356 TGCAGACATCTCAGTACGTGTCATCCCAAGCGGCGGCGCACTGCGGTGGCTTGA 1415
Db 1206 GCGCGACATCTGTGTAGTGTGATCCCTTACCGAGCTCCACCGCGAGGCTTCCA 1265
QY 1416 GAGCTTACCTCCCACTGCTGTGGGAGCGCGGAGCGCCCTACCTCATTTGGCTTTATCTC 1475
Db 1266 GATCGTGTCTCCCACTGCTGGGTGATGCTGGGAGCGCCCTACCTCATTTGGCTGTATCTC 1325
QY 1476 AGACCTGATCGCAGAGACTAAGGACTCCCGCTCTGGGAGTTCCTTGAGCTTGGGCTA 1535
Db 1326 TGACCGCTGCGCGGAACTGGCCCCCTCTCTTGTCCGAGTTCGCGGCTCTGAGTT 1385
QY 1536 GCGGCTCATGCTCTGCCCTTTCTGTCGTCTCTGGCGGCACTGTCTCTCTCGCCACTGC 1595
Db 1386 CTCGCTCATGCTCTCGGCTTTGTTGGGCACTGGGCGGCGGAGCTTCTCTGGCACCCG 1445
QY 1596 GCTCTTCTCTGTCAGCAGCCGCGCAGGGCTGAGCAGACCTTGGGG 1642
Db 1446 CATCTTCATTGAGGCGGACCGCGCGGCGGACAGCTGCAGCTGCAG 1492
```

```
RESULT 4
US-10-154-419-56
; Sequence 56, Application US/10154419
; Publication No. US20030143675A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Gluckman, Maria E.
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8099, 46455, 54414, 53763,
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FBH58295FL, 57255,
; TITLE OF INVENTION: AND 57255alt MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-249
; CURRENT APPLICATION NUMBER: US/10/154,419
```

; CURRENT FILING DATE: 2002-05-22  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 99  
; SOFTWARE: FastSeq Version 4.0  
; SEQ ID NO 56  
; LENGTH: 1587  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1587)  
US-10-154-419-56

Query Match 27.2%; Score 563.8; DB 15; Length 1587;

Best Local Similarity 64.3%; Pred. No. 5.6e-133;  
Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;

QY	282	GGGCGCGGCTCAGCAGCCCAACCGCCAGCTTGGCGCGCGGGGGGGGAGCGCGC	341
DB	126	GGACAGAGGGGCTGCAGCGCATCACCGGCTGTCTCCGGCGTTTCGGCTCTCATAGT	185
QY	342	CGGCATCTCAGCTTGGGCAACGTCTCACTACGACAGGTACACCGTGGCAGCGT	401
DB	186	GGCGGTCTGTCTACATCATCTCTGAACTACATGACCGCTTACCGTGGCTGGCGT	245
QY	402	CTTCTGGACATCCAGCAGCACTTTGGGGTCAAGGACCGAGCGCGCCCTGTCTGCAGTC	461
DB	246	CTTCCCGACATCGAGCAGTCTTCAACATCGGGGACAGTAGCTCTGGGCTCATCCAGAC	305
QY	462	AGTGTTCATCTGAGCTTATGTGGCTGCCCCATCTCGGTACCTGGGCGACCGCTT	521
DB	306	CGTGTTCATCTCAGTTACATGTGTGGCACCTGTGTGGCTTACCTGGGTGACAGGA	365
QY	522	CAACAGGAAGTGATCTCAGCTGCGCATTTCTCTGTGTCGCGCTCACCTTCTCCAG	581
DB	366	CAATCGAAGTATCTCATGTGGGGGCACTGCTCTGTGCTCCGTGTGACACTGGGGT	425
QY	582	CTCCTTCATTTCCAGCAGTACTTCTGGCTGTGCTCTCTGTCGGGGGCTGTGGGCAT	641
DB	426	ATCCTTCATCCCGGAGAGCATTTCTGGCTGCTCTCTGACCGGGGCTGTGGGGT	485
QY	642	CGGGGAGGCGAGTACTTCAACATCGCCCGCACTATCATTTGGGCACTCTTACCAAGAA	701
DB	486	CGGGGAGGCGAGTATTTCAACATCGCCCGCACTCTCATTTGGGCACTCTTTTGGCCGA	545
QY	702	CAGCGTACGCTCATGCTGTCGCTTCTTACTTCTGCTCCATCCACTGGGAGTGGCTGG	761
DB	546	CCAGCGAGCGGATGCTCAGCATCTTCTACTTTGCCATTCGGTGGGCGAGTGGTCTGG	605
QY	762	CTACATTTACTGGCTCCAGCGTGAAGCAGGCGAGGAGCTGGCACTGGGCACTTGGGGT	821
DB	606	CTACATTTGCAAGGCTCCAAAGTGAAGATATGGCTGGAGACTGGCACTGGGCTTGAGGT	665
QY	822	GTCCCTGTCTGGGCAATGATCAAGAACTCATCTCTATCTGTGTCCAGCCACTAA	881
DB	666	GACACCGGCTTAGGAGTGGTGGCGTCTGCTGCTGCTTCTGCTAGTGGGAGCGCC	725
QY	882	AAGGGTCTACGCGACAGCTCGGGA---CCAGCTCAAGCCCGGACCTCATGGCTCCG	938
DB	726	AAGGGAGCGTGGAGCGCCACTCAGATTTGGCACCCCTGAAACCCACCTCGTGGTGGG	785
QY	939	AGATATGAAGGCGCTGATTTGAAACCGCAGTACGTCTTCTCTCTCCCTGGCCACGTCGGC	998
DB	786	AGATCTGAGGCTCTGGCAAGAAATCTAGTTTCTGCTCTGCTCTCTCTGGCTTCTAGTC	845
QY	999	TGCTCTCTTCCGACCGGGGCGCTTGGGCAATGGATCCCGCTCTACTGACCGCGCCCA	1058
DB	846	TGTGGCTTTGTCAACGGGCTCCCTGGCTCTGTGGGCTCGGCATTCCTGTGGTTCGCG	905
QY	1059	AGTTGTGCA---GAAGACAGCAGAGCTGCACAGCCCGCTGTGGGCGCAAGACAG	1115
DB	906	CGTGGTCTTGGGGAGACCCACCTCTGCTTCCCGGAGACTCTCTCTCTCTCTGACAG	965

QY	1116	CCTCATCTTTGGGCGCATCATCTGCTTTTACGGGATTTCTGGCGTGTGTACGGGGGACGG	1175
DB	966	TCTCATCTTTGGGACTCATACCTGCTGACCGGAGTCTTGGGTGTGGGCTGGGTGTGGA	1025
QY	1176	AGCCACGGCTGGTGGCGCTGAAGACCCAGCGGGCCGACCCACTGGTGTGTGGCGTGGG	1235
DB	1026	GATCAGCGCGGCTCGGCCACTTCAACACCCCGGGCTGATCCCTGGTCTGTGGCACATGG	1085
QY	1236	CATGCTGGGCTCTGCCATCTTCTGCTGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1295
DB	1086	CCTCCTGGGCTCTGACCCCTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1145
QY	1296	AGGAGCCTATATCTGATCTTCTGCTGCGGGAGACGCTGCTGCTTCTTAACTGGGCGCATC	1355
DB	1146	GGCCACTTATATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1205
QY	1356	TGACAGCATCTCATGTAGTGTGTCATCCCAAGCGGGGCGGCACTGCTGCTGGGCTGGCA	1415
DB	1206	GGCGGACATTTCTGTGTACGTGTGATCCCTTACCCGACGCTCCACCGCGAGGCTTCCA	1265
QY	1416	GAGCTTCACTCCACCTGCTGGGGAGCGCGGGAGCCCTTCTTCTTCTTCTTCTTCTTCTT	1475
DB	1266	GATGCTGCTTCCACCTGCTGGGTGATGCTGGGAGCCCTTCTTCTTCTTCTTCTTCTTCT	1325
QY	1476	AGACCTGATCCGCGAGACACTAAGGACTCCCGCTCTGGGAGTTCCTGAGCCTGGGCTA	1535
DB	1326	TGACCGGCTCGCGCGGAACCTGCGCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1385
QY	1536	CGGCTCATGCTCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1595
DB	1386	CTCGCTCATGCTCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1445
QY	1596	GCTCTTCTTCTGCTCAGCGACCGCGCGGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGG	1642
DB	1446	CATCTTCTTCTGAGGCGGACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1492

## RESULT 5

US-10-146-733-51  
; Sequence 51, Application US/10146733  
; Publication No. US20030165891A1  
; GENERAL INFORMATION:  
; APPLICANT: Glucksmann, Maria A.  
; APPLICANT: Silos-Santiago, Immaculada  
; TITLE OF INVENTION: NOVEL TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615,  
; TITLE OF INVENTION: HNMDA-1, TWIK-9, alpha2delta-4, 54414, AND 53763 MOLECULES  
; TITLE OF INVENTION: AND USES THEREFOR  
; FILE REFERENCE: MNI-248  
; CURRENT APPLICATION NUMBER: US/10/146,733  
; CURRENT FILING DATE: 2002-05-15  
; PRIOR APPLICATION NUMBER: US 60/185,938  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 09/515,520  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 09/518,866  
; PRIOR FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: US 60/195,734  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: US 60/195,993  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: US 60/199,799  
; PRIOR FILING DATE: 2000-04-26  
; PRIOR APPLICATION NUMBER: US 60/233,537  
; PRIOR FILING DATE: 2000-09-19  
; PRIOR APPLICATION NUMBER: US 60/235,018  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US 60/235,059  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US 60/256,240  
; PRIOR FILING DATE: 2000-12-15  
; PRIOR APPLICATION NUMBER: US 60/256,588  
; PRIOR FILING DATE: 2000-12-18





US-10-648-593-127

Query Match 27.2%; Score 563.8; DB 17; Length 2165;  
Best Local Similarity 64.3%; Pred. No. 5.8e-133;  
Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;

```
QY 282 GGGCCCGGCGCTCAGACGCCAACAACCGGCCAGCTTGGGCGCGGGGGGGGGGAGCGGC 341
D 450 GGACAGAGAGGGGTGACAGCGCATCAACCGGCTGTCTCCCGGCGGTTCGGCTCTCATAGT 509
QY 342 CGCCATCTCTAGCTTGGGCAAGCTGCTCAACTACTGGACAGGTACACCGTGGCAGCGGT 401
D 510 GCGGCTGTGTCTACATCAATCTCTGAATCAATGACACCGCTTACCGTGGCTGGCGT 569
QY 402 CTTCTTGACATCCAGCAGCACTTTGGGGTCAAGGACGAGCGCGCCCTGTGTCAGTC 461
D 570 CTTTCCGACATCGAGCAGTTCTTCAACATCGGGGACAGTAGCTCTGGGCTCATCCAGAC 629
QY 462 AGTGTTCATCTGTAGCTTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 521
D 630 CGTGTTCATCTCCAGTTACATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 689
QY 522 CAACAGAGAGGTATCTCAGCTCGGCGATTTCTTCTGTGTCGGCGCTCACTTCTCCAG 581
D 690 CAATCGGAAGTATCTCATGTGCGGGGCAATGCTTCTGTGTCCTGTGTGACACTGGGGTC 749
QY 582 CTCCTTCAATTCGCCAGCAGTACTTCTGTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 641
D 750 ATCTTTCATCCCGGAGAGCATTTCTGTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 809
QY 642 CGGGGAGGCGAGTACTTCCACCATCGGCCGCCCACTATATGTTGGCACTCTTTCACCAAGAA 701
D 810 CGGGGAGGCGAGTATTTCCACCATCGGCCGCCCACTCTCATTTGCCGACCTCTTTTGGGCGGA 869
QY 702 CAGCGTACGCTCATGTCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 761
D 870 CCAGCGAGCGGATGTCTAGCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 929
QY 762 CTACATTACTGGCTTCCAGCGTGAAGCAGGAGCGCGAGACTGGCACTGGGCAATTCGGGGT 821
D 930 CTACATTGCAAGGCTTCCAAAGTGAAGATATGGCTGGAGACTGGCACTGGGCTCTGAGGGT 989
QY 822 GTCCCTCTGTCTGGGCAATATCAAGAACTCATCTCTATCTGTGTCAGCCAGCACTAA 881
D 990 GACACCGGGTCTAGGAGTGTGGCGCTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1049
QY 882 AAGGGTCTATGCCAGCAGCTCGGGG---CCAGCTCAAGCGCGGACCTCATGGCTCCG 938
D 1050 AAGGGGAGCGCTGGAGCGCCACTCAGATTTGCCACCCCTGAACCCCACTCTGTGTGGGC 1109
QY 939 AGATATGAAGCGCCCTGATTGMAACCGCAGCTAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCT 998
D 1110 AGATCTAGGGCTCTGCAAGAAATCTAGTTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1169
QY 999 TGTCTCTCTTCCGCAAGGGGCGCTTGGGCAATGTGGATCCCGCTCTTACCTGCAACCGCGCCA 1058
D 1170 TGTGGCTTGTCTACGGGCTCCCTGTCTGTGGGCTCCGCACTTCTCTGTCTGTCTGTCTGTCT 1229
QY 1059 AGTTGTGCA---GNAAGACAGAGAGAGCTGCAACAGCGCGCGCTGTGGGCGCAAGAGCAG 1115
D 1230 CGTGGTCTTTGGGGAGAGCCCACTCCCTTCCCGGAGACTCTCTGTCTTCTCTCTCTCTCTCT 1289
QY 1116 CCTCATCTTTGGGCGCATCACTGTCTTTACGGGATTTCTGGGCTGTGTGTGTGTGTGTGTGTGT 1175
D 1290 TCTCATCTTTGGACTCATCTGCTGTGACGGAGTCTCTGGGTGTGGGCTGTGGTGTGGA 1349
QY 1176 AGCCAGCGCTGTGGCGCTGGAAGACCCAGCGGCGGACCCCACTGTGTGTGTGTGTGTGTGTGT 1235
D 1350 GATCAGCGCGCGGCTCGCGCACTCCAACCCCGGGCTGATCCCTGTGTGTGTGTGTGTGTGTGT 1409
QY 1236 CATGCTGGGCTCTGCCATCTTCTATCTGCTGTATCTGTGGTGTGCCAAGAGCAGCATCGT 1295
D 1410 CTTCTGGGCTCTGCAACCTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1469
```

```
QY 1296 AGGAGCCTATATCTGTATCTTGTGGGGAGACGCTGTCTTTTCTTAACCTGGGCATCAC 1355
D 1470 GGCCACTTATATTTTCTATCTTCTTATGAGAGAGACCTCTCTGTCATGAACCTGGGCATCGT 1529
QY 1356 TGCAGACATCTCATGTACGTGTGTCATCCCAACGCGGGGCGGCACTGCGGCTGGCTTGA 1415
D 1530 GCGCGACATTTCTGTGTACGTGTGATCCCTACCCGACGCTCCACCGCGAGGCTTTCCA 1589
QY 1416 GAGCTTCACTCTCCACCTGTGGGGGACGCGGAGCCCTTACCTCATTTGGCTTTATCTC 1475
D 1590 GATGCTGCTCTCCACCTGTGGGTGATGTGGAGGCCCCCTACCTCATTTGGCTGTATCTC 1649
QY 1476 AGACCTGATCCGCGAGACACTAAGGACTCCCGCTCTGGGAGTTCTTGAGCCTGGGCTA 1535
D 1650 TGACCGCTGTGGCGGAACTGGCCCCCTCTCTTCTGTGCGAGTTCCGGGCTCTGCAGTT 1709
QY 1536 GCGGCTCATGCTGTGCGCTTTCGTGCTGTGCTGTGGGGGCGCATGTTCTCTGCGCACTGC 1595
D 1710 CTCGCTCATGCTGTGCGCTTTCGTGCGGCACTGGGCGGCGGAGCTTCTTGGGCGACCGC 1769
QY 1596 GCTCTTCTCTGTCAGCGACCGCGCCAGGCTGTGAGCAGCACTGGGGG 1642
D 1770 CATCTTCAATGAGGCCCGACCGCGCGGCGACAGCTGCACGTGCGAG 1816
```

RESULT 7

```
US-10-098-841-56
; Sequence 56, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 56
; LENGTH: 2218
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (372)..(1958)
US-10-098-841-56
```

Query Match 27.2%; Score 563.8; DB 13; Length 2218;  
Best Local Similarity 64.3%; Pred. No. 5.8e-133;  
Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;







861	CGGGAGGCCAGTTATTATCCACATCGCGCCCACTCTCATTTGCCAGACTCTTTTGTGGCCGA	920
702	CACCGTACGCTCATGTCTGTCGGTCTTCTACTTCGCCCATCCCATCGGCGAGTGGCCTGGG	761
921	CCAGCGAGCCGATGCTCAGCATCTTCTACTTTTGCCATTCCGGTGGCGAGTGGTCTGGG	980
762	CTACATTTACTGGCTCCAGCGTGAAGCAGGACGCCGGAGACTGGCACTGGGGCATTTGCCGGT	821
981	CTACATTTGCAGGGCTCCAAAGTGAAGGATATATGGCTGAGAGACTGGCACTGGGCTCTGAGGGT	1040
822	GTCCCTCTGTCTGGCGCATGATCACAGGAACACTCATCTCATTTCTGTGTCACAGCCACTAA	881
1041	GACACCGGCTTAGGAGTGGTGGCGCTTCTGCTGCTGTTCTGTGTAGTCGGGAGCCGCC	1100
882	AAGGGGTATGTCGCGACAGCTCGGGGA---CCAGCTCAAGGCCCGGACCTCATGGCTCCG	938
1101	AAGGGGAGCGTGGAGCGCACTCAGATTTGCCACCCCTGAACCCCACTCGTGGTGGGC	1160
939	AGATATGAAGGCCCTGATTTGAAACCGCAGGTAGTCTTCTCTCCCTCGGCACAGTCGGC	998
1161	AGATCTGAGGGCTCTGGCAAGAAATCCTAGTATTGCTCTGTCTTCTCCCTGGGCTTCACTGC	1220
999	TGTCCTCTTCGCCACCGGGGCCCTGGGCATGTGATCCCGCTCTACTCTGCACCGCGCCA	1058
1221	TGTGGCCCTTTGTACGGGCTCCCTTGGCTCTGTGGGCTCCGGCATTCCTGCTGGTTCGGC	1280
1059	AGTTGTGTGA---GAAGACAGCAGAGAGCTGCACACAGCCCGCCCTGTGGGGCCCAAGACAG	1115
1281	CGTGGTCTTTGGGAGACCCACCCCTGCCCTTCCGGAGACTCTGTCTTCTCTGACAG	1340
1116	CCTCATCTTTGGGGCCATGACCTGTCTTACGGGATTTCTGGGCGTGGTCAAGGGGGCAGG	1175
1341	TCCTCATTTTGGACTCATACCTGCTCCAGCCGAGTCTCGGGTGTGGGCTCTGGTGTGGA	1400
1176	AGCCACGCGCTGTGTCCGCTGAAAGACCCACAGCGGGCCGACCACTGGTGTGTGCCGTGGG	1235
1401	GATCAGCGCGGCTCGCGCACTCCAAACCCCGGGCTGATCCCTGGTCTGTGCGCACTGG	1460
1236	CATGCTGGGCTCTGCGCATTTCTATCTGCTGATCTTGTGGTGTGCCAAGACAGACATCGT	1295
1461	CCTCTGGGCTCTGACCCCTCTCTTCTCTGTCCCTGTGCCCTGGCGCCGTGGTAGCATCGT	1520
1296	AGGAGCTATATCTGTATCTTCGTGGGAGACGCTGCTGTTTTTCTAACTGGGGCCATCAC	1355
1521	GGCCACTTATATTTTCACTTTCATTTGAGAGACCCCTCTGTCATGAACCTGGGCCATCGT	1580
1356	TGCAGACATCCTCATGTACGTGTGTCATCCCAAGCGCGGGCCACTGCGCGTGGCTTGCA	1415
1581	GGCCGACATTTCTGTGTACGTGTGATCCCTACCCAGAGCTCCACGCCCGAGGCTTCCA	1640
1416	GAGCTTCACTCTCCCACTGTCTGGGGAGCGCGGGAGCCCTACTCTCATTTGGCTTTATCTC	1475
1641	GATCGTGTCTGCCCACTGTGGGTGATGCTGGGAGCCCTACTCTCATTTGGCCTGATCTC	1700
1476	AGACCTGATCCGCCAGAGCACTAAGGACTCCCGCTCTGGGAGTTCCTGAGCCTGGGCTA	1535
1701	TGACCGGCTCGGCGGAACCTGGCCCCCTCTTCTTGTCCGAGTTCGGGCTCTGCAGTT	1760
1536	CGGCTCATGCTCTGCCCTTTTCGTGTGTGCTCTGGGCGGCAATGTTCTTCTCGCCACTGC	1595
1761	CTCGCTCATGCTCTGGCGGTTTGTGGGGCACTGGGGGGCGAGCTTCTCTGGGACCGC	1820
1596	GCTCTTCTTCGTGAGCGACCGCGCCAGGGCTGAGCAGCACTGGGGG	1642
1821	CATCTTTCATTGAGGCCGACCGCGGGGGGACAGCTGCAGTGCAGG	1867

## RESULT 11

US-09-833-381-883

03-03-033-381-803  
: Sequence 883. Application US/09833381

; sequence 883, APPLICATION  
: Patent No. US20020132090A1

; PATENT NO. US2002013

APPLICANT: Robison, Keith E.





Db 614 GATCGTGTGTCACACCTGCTGGGTGATGCTGGGAGCCCTACCTCATTTGGCCGTGATCTC 555  
Qy 1476 AGACCTGATCCGCGAGACACTAAGGACTCCCCCGCTCTGGGAGTTCTCTGAGCCTGGGCTA 1535  
Db 554 TGACCGCTGCGCGGAAGTGGCCCTCTCTTCTTGTCCGAGTTCCGGGCTCTGCAGTT 495  
Qy 1536 CGCGCTCATGCTCTGCCCTTTCTGCTGTGCTGCTGCTGGGCGGATGTTCTTCTGCGCACTGC 1595  
Db 494 CTCGCTCATGCTCTGCGCGTTTGTGTGGGCACTGGGCGGCGCAGCYTTCTTGGGCACGCG 435  
Qy 1596 GCTCTTCTCTGATGACCGCGCCAGGGCTGAGCAGCACCTGGGG 1642  
Db 434 CATCTTCAATGAGCGCCAGCCCGCGGGGACAGCTGCAGTGCAAG 388

## RESULT 13

US-09-809-391-245  
; Sequence 245, Application US/09809391  
; Publication No. US20030049618A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P2  
; CURRENT APPLICATION NUMBER: US/09/809,391  
; CURRENT FILING DATE: 2001-03-16  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 761  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 245  
; LENGTH: 2201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-809-391-245

Query Match 26.28; Score 542.8; DB 10; Length 2201;

Best Local Similarity 64.38; Pred. No. 1.2e-127; Mismatches 478; Indels 8; Gaps 4;

Matches 877; Conservative 1; Mismatches 478; Indels 8; Gaps 4;

Qy 285 CCCCGCGCTCAGCAGCCAAACCGCGCAGCTTGGGCGCGGGCGGGGCGAGCCGCGC 344  
Db 416 CCAGGARGGCTGCAGCGCATCMCGGCTGTCTCCGCGCGTTCGGCTCTCATATGGC 475  
Qy 345 CATCTCAGCTTGGGCAACGTGTCTCAACTACTGGACAGTACACCGTGGCAGCGCTCT 404  
Db 476 GGTGCTGTGTACATCAATCTCTGAACTACATGACCGCTTCAACGCTGCTGGCGTCT 535  
Qy 405 TCTGGACATCCAGCAGCACTTGGGTCAAGGACCGAGCGCGCGCTGTGAGTCAGT 464  
Db 536 TCCGACATCGAGGAGTCTTCAACATCGGGGACAGTAGCTCTGGGCTCATCCAGACCGT 595  
Qy 465 GTTCATCTGTAGCTTTCATGTGGCTGCCCATCTTCCGCTACTCTGGGCGACCGCTTCAA 524  
Db 596 GTTCATCTCAGTTACATGTGTGGACCTGTGTGGTGTGCTGTGGTGTGAGTACAGTACAA 655  
Qy 525 CAGGAAGTGATTTCTCAGTCTCGGCAATTTCTTCTGTGTGGCGCTCACTTCTCCAGCTC 584  
Db 656 TCGGAAGTATCTCATGTGCGGGGCAATTCCTTCTGTCTCTGTGTGACACTGGGGTCACTC 715  
Qy 585 CTTCATTTCCCGCAGTACTTCTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 644  
Db 716 CTTCATTTCCCGGAGAGCAATTTCTGGCTGTCTCTCTGACCGCGGGGCTGTGTGGGGTCTG 775  
Qy 645 GGAGGCGAGTACTCTCACCATCCGCCCACTATCATTTGGGACCTCTTCAACCAAGAACAC 704  
Db 776 GGAGGCGAGTATTTCAACCATCCGCCCACTCTCATTTGGGACCTCTTGTGGCGGACCA 835  
Qy 705 GGTACGCTCATGCTGTCCGTCTTCTTACTTGGCATCCCACTGGGAGTGGCTGGGCTA 764  
Db 836 GCGGAC-CGGATGCTCAGCATCTTCTTACTTTGGCATTCGGGTGGGAGTGGTCTGGGCTA 894  
Qy 765 CATTACTGGCTCAGCGTGNAGCAGGCGGAGACTGGCACTGGGCACTGGCGGTGTC 824  
Db 895 CATTGCGGCTCCAAAGTGAAGATATGGCTGGAGACTGGCACTGGGCTCTGAGGGGTGAC 954

Qy 825 CCCTGTCTCTGGGCAATGATCAGGAACACTCATCTCTATTCTGTGCTCCAGCCACTTAAAG 884  
Db 955 ACCGGTCTAGGAGTGTGGCGTTCGTCTGTCTGTCTGTGTAGTGGGAGCCGCCAAG 1014  
Qy 885 GGTGATGCGGACAGCTCGGGG---CAGCTCAAGCCCGGACCTCATGGCTCCGAGA 941  
Db 1015 GGGAGCGGTGGAGCGCCACTCAGATTTGCCACCCCTGAACCCACCTCGTGTGGGAGA 1074  
Qy 942 TATGAAGCGCTGATTTCGAACCGCAGCTAGTCTTCTCTCCCTGGCCACGTCGGCTGT 1001  
Db 1075 TCTGAGGGCTCTGGCAAGAAATCCTAGTTTCTGTCTTCTCTCTGGCTTCACTGCTGT 1134  
Qy 1002 CTCCTTCCGACCGGGGCTCTGGGCAATGTGATCCCGCTCTACCTGACCCCGCCCAAGT 1061  
Db 1135 GGCCTTTGTCAAGCGCTCTCTGGCTCTGTGGGCTCCGCAATCTCTGCTGCTTCCGCGT 1194  
Qy 1062 TGTGCA---GAAAGACAGAGAGAGTGAACAGCCCGCCCTGTGGGGGCAAGAGCAGCT 1118  
Db 1195 GGTCTTTGGGAGAGACCCCTGCTTCCCGGAGACTCTCTGCTTCTCTGACAGTCT 1254  
Qy 1119 CATCTTTGGGCGCATCACCCTGCTTACCGGATTTCTGGGCGTGTGTCAGGGGGCAGGAGC 1178  
Db 1255 CATCTTTGGACTCATCACCCTGCTGACCGGAGTCTGGGTGTGGGCTGTGGAGAT 1314  
Qy 1179 CAGCGCTGTGCGCGCTGAAAGACCCAGCGGCGGACCCACTGTGTGTGTGCTGGGCGAT 1238  
Db 1315 CAGCGCGGCTCGGCACTCTCAACCCCGGGGTGATCCCTGTGTGTGCTGCTGCTGGCT 1374  
Qy 1239 GCTGGGCTGTGCATCTTCACTGTCTGTGATCTTGTGTGGCTGCCAAGAGCAGCATCTAGG 1298  
Db 1375 CTTGGGCTGTGACCCCTTCTCTTCTGCTGCTTCTGCTGCGCCGCTGTAGCATCTGGC 1434  
Qy 1299 AGCTATATCTGTATCTTCTCGGGGAGAGCTGCTGTCTTCTTACTTGGGCGCATCTGC 1358  
Db 1435 CACTTATATTTTATCTTCAATGGAGAGACCTCTCTGTCCATGAATCTGGGCGCATCTGC 1494  
Qy 1359 AGACATCTCATGTACGTGTGTATCCCAAGCGGCGGCGCACTGCCGTGGCTTGTGACAG 1418  
Db 1495 CGACATCTCTGTGATGTGTGTATCTCCACAGCTTCCACCGCCAGGCTTCCAGAT 1554  
Qy 1419 CTTCACCTCCACTGTCTGGGAGACCGCGGAGCCCTACTCATTTGGCTTTATCTCAGA 1478  
Db 1555 CGTGTCTGCCACTGTCTGGGTGATGTGGGAGCCCTACTCATTTGGCTGTATCTCTGA 1614  
Qy 1479 CTTGATCCGCGCAGCACTAAGCACTCCCGCTCTGGGAGTTCCTGAGCCTGGGCTAGC 1538  
Db 1615 CCGCCTGCGCGGAACTGGGCCCTCTTCTTCTCGAGTTCGGGCTCTGCAAGTTCTC 1674  
Qy 1539 GCTCATGCTCTGCCCTTTCTGTGTGTGCTGTGGGCGCATGTTCTTCTCTGCCACTGCGCT 1598  
Db 1675 GCTCATGCTCTGGCGTTTGTGGGCACTGGGGGGGCGCACT-TTCTTGGGACCGGCAAT 1733  
Qy 1599 CTTCTTCTGTCAGCGACCGCGCCAGGGCTGAGCAGCACTGGGGG 1642  
Db 1734 CTTTATTGAGCGGACCGCGCGGGGCGACAGCTGACGTCAGG 1777

## RESULT 14

US-09-882-171-245  
; Sequence 245, Application US/09882171  
; Publication No. US20030175858A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P2  
; CURRENT APPLICATION NUMBER: US/09/882,171  
; CURRENT FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 09/809,391  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 09/149,476  
; PRIOR FILING DATE: 1998-09-08  
; PRIOR APPLICATION NUMBER: PCT/US98/04493

9





Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	666	32.2	668	4	BM722582	UI-B-E00-	
2	646.8	31.2	958	4	B1410266	B1410266	
3	630.8	30.5	924	4	B1414356	B1414356	
4	596.6	28.8	828	7	CN528616	UI-N-HQ0-	
5	575.8	27.8	772	7	CK479120	AGENCOURT	
6	557.8	26.9	837	4	B1159625	602920091	
7	552.8	26.7	2144	3	AF370423	Homo sapi	
8	524.2	25.3	543	4	BM147219	TCARP1Q10	
9	512.4	24.7	1075	2	B285645	BM147219	
10	509.8	24.6	687	4	BG243478	B285645	
11	506.6	24.5	524	7	CR556576	602356937	
12	502.6	24.3	815	7	CO401999	DKFZp459J	
13	498.8	24.1	724	4	BG974364	AGENCOURT	
14	495.2	23.9	1587	9	AY403401	602844132	
15	483	23.3	579	7	CN666419	AY403401	
16	474.6	22.9	1587	9	AY403403	Homo sapi	
17	471.2	22.8	549	2	B233986	A0840B03-	
18	470.8	22.7	1430	3	CR591379	Mus muscu	
19	470	22.7	1615	3	CR596645	140546 MA	
20	468.8	22.6	771	2	BF099708	full-len	
21	467.4	22.6	989	4	BG176211	CR596645	
22	453	21.9	518	5	BX280500	full-len	
23	442.6	21.4	668	2	B377721	CR596645	
24	442	21.3	840	4	BI765222	BF099708	
25	442	21.3	840	4	BI765222	601751343	
26	442	21.3	840	4	BI765222	60235356	
27	442	21.3	840	4	BI765222	BX280500	
28	442	21.3	840	4	BI765222	601229867	
29	442	21.3	840	4	BI765222	601229867	
30	442	21.3	840	4	BI765222	603050307	

```

FEATURES
    seq primer: M13 Reverse.
    Location/Qualifiers
        1..668
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="UI-E-E00-ahy-f-02-0-UI"
            /tissue_type="fetal eye"
            /dev_stages="fetal"
            /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
            /clone_lib="UI-E-E00"
            /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-E00 is a cDNA library containing the following
tissue(s): fetal eye. The library was constructed

```



Db 366 TGGTGGCCCTCGGAGACTCAACGCTGCTACCCATGGGTGTGTGTGGGATCGTGGGA 425  
Qy 1246 TCTGCCATCTTCATCTGCTGATCTTCGTGGCTGCCAAGACAGACATCGTAGGAGCCTAT 1305  
Db 426 TCCGCCATCTTCATCTGCTCATCTTGTGGCTGCCAAGACAGACATGTAGGGGCTTAT 485  
Qy 1306 ATCTGTATCTTCGTGGGAGACGCTGCTGTTTTCTAACTGGGCACTACATGACAGATC 1365  
Db 486 ATCTGCAATTTTGTGGAGAGACATCTGCTGTTTTCTAACTGGGCACTACATGACAGATC 545  
Qy 1366 CTCATGTACGTGGTCAATCCCAACGCGCGCCACTCGCGTGGCTTGCAGAGCTTCACC 1425  
Db 546 CTCATGTATGGGTCAATCCCAACGCGCGCCACTCGCGTGGCTTGCAGAGCTTCACC 605  
Qy 1426 TCCCACTGTCTGGGAGCGCGGAGCGCCCTACCTCATTTGGCTTTATCTCAGACCTGATC 1485  
Db 606 TCCCACTGTCTGGGAGCGCTGGAAGCCCTACCTCATTTGGCTTTATCTCGGACTAATC 665  
Qy 1486 CGCAGAGCACTAAGGACTCCCGCTCTGGAGTTCTTGAGCCTGGGCTAGCGCTCATG 1545  
Db 666 CGCAGAGCACTAAGGACTCCCGCTCTGGAGTTCTTGAGCCTGGGCTAGCGCTCATG 725  
Qy 1546 CTCTGCCCTTTCGTGGTGGTCTCGGCGGCA-TGTTCTTCTCGCACTGGGCTTCTT 1604  
Db 726 CTGTGCCCTTTCGTGGTGGTCTCGGCGGCA-TGTTCTTCTCGCACTGGGCTTCTT 785  
Qy 1605 CGTCAGGACCGCGCCAGGCTGAGCAGCACTCTGGGAGAGACGGCGGGGTCAGGGT 1664  
Db 786 CCTCAGGACCGCTGCCA-GGCTGAGCAGCAGGTGAACACAGCTGTGTATGCTCTCGCATCC 844  
Qy 1665 GGTGCATCAGCGGGCGCGGCGCGGCACTGCT 1698  
Db 845 GGAAAGTCTCAGGGGTGCGGCTGGGACAGCACT 878

RESULT 3  
BI414356 924 bp mRNA linear EST 14-AUG-2001  
LOCUS 602990376F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5146410 5',  
DEFINITION mRNA sequence.  
ACCESSION BI414356  
VERSION BI414356.1 GI:15175279  
KEYWORDS Mus musculus (house mouse)  
SOURCE EST.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 924)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-x@mail.nih.gov](mailto:cgapbs-x@mail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM11361 row: 9 column: 19  
High quality sequence start: 10  
High quality sequence stop: 824.  
Location/Qualifiers  
1. .924  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CZECH II"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5146410"  
/tissue\_type="pooled lung tumors"

FEATURES  
source

/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NCI\_CGAP\_Lu33"  
/note="Organ: lung; Vector: p773D-Pac (Pharmacia) with a  
modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st  
strand cDNA was prepared from mRNA obtained from pooled  
lung tumors with a Not I - oligo(dT) primer [5'.  
TGTTACCAATCTGAAGTGGAGCGCGCTCTGTTTTTTTTTTTTTTT 3'].  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not  
I and Eco RI sites of the modified p773 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 30.5%; Score 630.8; DB 4; Length 924;  
Best Local Similarity 86.5%; Pred. No. 8.4e-118;  
Matches 754; Conservative 0; Mismatches 112; Indels 6; Gaps 5;  
Qy 755 GCCTGGGCTACATTACTTGGCT--CCAGCGTGAAGCAGCGAGCTGGGAGACTGGGC 812  
Db 8 GGCTGGGCTATATCACAGGCTTCAGCGTTGAAGCAGCGAGCTGGGAGCTGGGC 67  
Qy 813 ATTCGGGCTCTCCCTGCTCTGGGCTATGATCAGAGGAACTCATCTCTCATCTTGTGCC 872  
Db 68 CCTGGGGTTTCCCGCTCTGGGCTATGATCAGAGGAACTCATCTCTCATCTTGTGCC 127  
Qy 873 AGCCACTAAAGGGGTCTATCCGACAGCTCGGGGACAGCTCAAGCCCGGACCTCATG 932  
Db 128 AGCCACTTAAGAGGCGCTATGATCACTTTGGGGGAGCTCAAGACAGGAGCTCTCTG 187  
Qy 933 GCTCCGAGATATGAAGGCCCTGATTGAAACCGCAGCTACGTCTTCTCTCTCTGGCCAC 992  
Db 188 GCTTCGAGACATGAAGGCCCTGATCCGAAACCGCAGTTAGCTTCTCTCTCTGGCCAC 247  
Qy 993 GTGGGCTGTCTCTTCCGACAGGGGGCTTGGGCTATGATCCGCTCTACCTGACCG 1052  
Db 248 ATCCGCTGTCTCTTCCGACAGGGGGCTTGGGCTATGATCCGCTCTATCTTCAACG 307  
Qy 1053 CGCCCAAGTTGTGCAGAGACAGCAGAGAGCTGCAACAGCCCGCTGTGGGGCCAGGA 1112  
Db 308 TGTCTAAGTTGTACAAAGACAGCAGAGACTTGCACAGCCCGCTTGTGGAGCAAGA 367  
Qy 1113 CAGCCTCATCTTTGGGCGCATCACTGCTTTACGGGATTTCTGGGCGTGTGTCAGGGGG 1172  
Db 368 CAGCCTCATCTTTGGGCGCATCACTGCTTTACTGGCTTCTGGGCGTAGTCACGGGCG 427  
Qy 1173 AGAGCCACCGCTGTGGCGCTGAAGACCCAGCGGCGGACCACTGGTGTGTCGGT 1232  
Db 428 GGGAGCCACTCGTTGGTGGCGCTTGGGACTCAACGCTGCTGAGCCATTGGTGTGCTGT 487  
Qy 1233 GGGCATGCTGGGCTCTGCCATCTTCATCTGCCCTGATCTTCTGGCTGCCAGAGCAGCAT 1292  
Db 488 GGGCATGCTGGGATCCGCCATCTTCATCTGCCCTGATCTTGTGGCTGCCAGACAGCAT 547  
Qy 1293 CGTAGGAGCCTATATCTGTATCTTCGTGGGAGAGCGCTGCTGTTTTCTAACTGGGCGCAT 1352  
Db 548 TGTAGGCGCTTATATCTGCAT-TTTGTGGAGAGACACTGCTGTTTTCTAACTGGGCGCAT 606  
Qy 1353 CACTGAGACATCTCATGTATGTCATCCACAGCGGCGGCGCACTGCGGTGGCGCTT 1412  
Db 607 CACTGAGACATCTCATGTATGTCATCCCACTCTGGGAGGCGCACTGCTGTGGCGCT 666  
Qy 1413 GCAGAGCTTCACTCTCCACCTGCTGGGAGCGCGGAGCGCCCTACCTCATCTGGCTTTAT 1472  
Db 667 GCAGAGCTTCACTCTCCATCTGCTGGGAGCGCTGGAAGCCCTTACCTCATTTGCG-TTAT 725  
Qy 1473 CTCAGACCTGATCCGCGCAGACACTAAGGACTCCCGCTCTGGGAGTTCTGAGCCCTGGG 1532  
Db 726 CTCGAGACCTAATCCG-CAGAGACCAAGGACTCCCGCTCTGGAGTCCCTGAGCCCTGGG 784  
Qy 1533 CTACGCGCTCATGCTCTGCCCTTTCTGCTGTGCTCTGGGCGGCAATG-TTCTTCTCGCCA 1591  
Db 785 CTATGCCCTCATGCTGTGCCCTTTTGTGCTGTGCTCTGGGCGGCAATGTTTCTTCTTGGCCA 844







|||||  
Db 259 GACCTCTGGCTTCGAGACATGAAGCCCTGTATCCGAAACCGCAGTACGTTCTCTCTC 318  
Qy 984 CTTGCCACGCTGGCTGTCTCTCGCCACGGGGCCCTGGGCATGTGGATCCCGCTCTA 1043  
Db 319 CTTGCCACATCGCTGTCTCTCGCCACAGGGCCCTGGGCATGTGGATCCCGCTCTA 378  
Qy 1044 CTTGCCACCGCCCAAGTTGTGCGAGAGACAGACAGAGAGCTGCAACAGCCCGCCCTGTGG 1103  
Db 379 TCTTCAACCGTCTCAAGTTGTACAAAGACAGACAGAGACTTGCAACAGCCCGCCCTGTGG 438  
Qy 1104 GCGCAAGGACAGCTCATCTTTGGGGCCATCACTGCTTTACGGGATTTCTGGGGTGTGT 1163  
Db 439 AGCCAAAGACAGCCCTCATCTTTGGAGCCATTACCTGCTTTACTGGCTTCTCGGGGTAGT 498  
Qy 1164 CACGGGGGACGAGGACACGCGCTGTGGCGCTGGAAGACCCAGCGGGCGGACCCCACTGGT 1223  
Db 499 CACGGGGGAGGAGCCACTCGTTGGTGGCGCTGCGGACTCAACGCTGACCCATTGGT 558  
Qy 1224 GTGTGCGGTGGGATGCTGGGCTCTGGCCATCTTCACTGCTGCTGATCTTGGTGGCCNA 1283  
Db 559 GTGTGCTGTGGGATGCTGGGATCGGCCATCTTCACTGCTGCTCATCTTTGTGGCTGCCAA 618  
Qy 1284 GAGGACATCGTAGAGCCATATATCTGTATCTTGTGCGGGAGAGCGCTGCTGTTTCTAA 1343  
Db 619 GACCAAGCATGTAGGGCGCTTATCTGCAAT-TTGTGTGAGAGACACTGCTGTTTCTTAA 677  
Qy 1344 CTTGGGCATCACTGCAGACATCTCATGTACGTGTGTCATCCCAACGCGCGGCCCACTGC 1403  
Db 678 CTTGGGCATCACTGCAGACATCTCATGTATGTGTGTCATCCCACTCGCGGAGCCACTGC 737  
Qy 1404 CGTGGCCTTGCAGAGCTTCACTTCCACCTGCTGGGGAGCGCGGGAGCCCTTACTCAT 1463  
Db 738 TGTGGACCTGCAGAGCTTCA-CTCCCATCTGCTGGGGAGCGGTGGAAAGCCCTTAACATCAT 796  
Qy 1464 TGGCTTTATCTAGACCTCATCGCCAGAGACACTAAGGACTCC 1506  
Db 797 TGG-TTTATCTGGACCTTATCCG-CAGAGCAAGGACTCC 837

RESULT 7  
AF370423 2144 bp mRNA linear HTC 31-DEC-2003  
LOCUS Homo sapiens PP2030 mRNA, complete cds.  
DEFINITION AF370423  
ACCESSION AF370423.1 GI:33341775  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2144)  
AUTHORS Huang, Y., Zhou, X.M., Zhang, P.P., Jiang, H.Q., Qin, W.X., Zhao, X.T.,  
Wan, D.F. and Gu, J.R.  
TITLE Novel human cDNA clones with function of inhibiting cancer cell  
growth  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2144)  
AUTHORS Zhou, X.M., Zhang, P.P., Jiang, H.Q., Huang, Y., Qin, W.X., Zhao, X.T.,  
Wan, D.F. and Gu, J.R.  
TITLE Direct Submission  
JOURNAL Submitted (16-APR-2001) National Laboratory For Oncogenes & Related  
Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai  
200032, P. R. China  
FEATURES  
source Location/Qualifiers  
1. .2144  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
597..1859  
/codon\_start=1  
/product="PP2030"  
/protein\_id="AAQ15259.1"  
CDS

/db\_xref="GI:33341776"  
/translation="MVLAPVRGYLDRNRYLMCGGIAPFWSLTVLGSFIPGHEFWL  
LLLRGLVGEASYSYSLPTLIADLVADORSMLSFYFAI PVGSLGLYAGSKVK  
DMAGDHWALRVTPGLGVAVLLFLVVRPGRGAVRHSDLPLNPFSWADLRALA  
RNPFSVLSSGFTAVFVSLRSLRPAFLRVRVLTGTPPCPLPGDSCSSDSLIFG  
LITCLTGLVGLGVEISRIRHSNPRADPLVATGLGSAFFLFLSLACARGSLVAT  
YLFITGISTLLMNNAIYADILLYVVIPTRRSTAEAFQIVLSHLLGAGSPYLGLIS  
DLRNNPPSPFLSEFRALQFLMLCAFVGLGAAFLGTAIFIBADRRRAQLHVQGLL  
HEAGTDDRIVVPRGRSTRVPASVLI"  
ORIGIN  
Query Match 26.7%; Score 552.8; DB 3; Length 2144;  
Best Local Similarity 64.3%; Pred. No. 6.5e-102;  
Matches 879; Conservative 0; Mismatches 482; Indels 7; Gaps 3;  
Qy 282 GGGCCCCGGCGCTCAGACGCCAAACCGGCAGCTTGGGCGCGGGCGGGGGGCGAGCCGC 341  
Db 397 GGACAGAGGGGCTGCAGCGCATCACCGGCTGTCTCCCGGCCCTTCGGGCTCTCATAGT 456  
Qy 342 CGCCATCTCAGCTTGGGCAACGTCTCAACTACCTGACAGGTATACACCTGGCAGGGCT 401  
Db 457 GCGGTGCTGTGTACATCATCTCTGAACTATACATGACCGCTTACCGTGGCTGGCT 516  
Qy 402 CTTCTGGACATCCAGC-AGCACTTTGGGGTCAAGGACCGAGCGCCGGCCTGTGCTGCT 460  
Db 517 CTTTCCCGACATCGAGCAAGTTCTTCAACATCGGGGACAGTAGCTCTGGGCTCATCCAGA 576  
Qy 461 CAGTGTTCATGTAGCTTCACTGTGGTGGTGGCCCATCTTCGGCTACCTGGGCGAGCCGT 520  
Db 577 CCGTGTTCATCTCCAGTTACATGGTGTGGCACCTGTGTTGGCTACCTGGGTGACAGT 636  
Qy 521 TCAACAGGAAGTGATTTCTCAGCTCGGCAATTTCTTCTGTGTCGGCGCTCACCTTCTCCA 580  
Db 637 ACATCGGAAGTATCTCATGTGCGGGGCAATGCTTCTGTGCTCCCTGGTGACACTGGGCT 696  
Qy 581 GCTCTTTCATTTCCCGAGCAGTACTTCTGGCTGCTGGTCTGTCTGTCGCGGGGCTGTGGGCA 640  
Db 697 CATCTTCATCCCGAGAGCAATTTCTGGCTGCTCTCTGACCCCGGGGCTGGTGGGG 756  
Qy 641 TGGGGAGGCGGCTACTCCACATCGCCCGCCACATATCATTTGGCGACCTCTTCCACAGA 700  
Db 757 TCGGGAGGCGGCTTATTTCCACATCGCGCCCATCTCATTTGCCGACCTCTTTGTGGCGG 816  
Qy 701 ACACGCGTACGCTCATGTGTCGCTTCTTACTTCGCCATCCCACTGGGCACTGGGCTGG 760  
Db 817 ACCAGGGAGCGGATGCTCAGCACTTCTTACTTTTGGCAATTCGGTGGGCACTGGTCTGG 876  
Qy 761 GCTACATTACTGGCTCCAGCGTGAAGCAGGAGCCGAGACTGGGCACTGGGCACTGGCGG 820  
Db 877 GCTACATTTGACGGCTCCAAGTGAAGGATATGGCTGGAGACTGGGCACTGGGCTCTGAGGG 936  
Qy 821 TGTCTCCCTGTCTCGGCGCATGATCAGAGGAACACTCATCTCATTTCTGGTCCGACCCACTA 880  
Db 937 TGACACCGGCTTAGAGTGTGGCGCTTCTGCTGTCTCTGTTCTGTGTAGTGGGAGCCCGC 996  
Qy 881 AAAGGGGTCAATCCGACCACTCGGGGA---CAAGCTCAAGGCCCGGACCTCATGCTGCTCC 937  
Db 997 CAAAGGGAGCCGTGGAGCGCCACTCAGATTTGGCAACCCCTGAACCCGACCTCGTGGTGGG 1056  
Qy 998 GAGATATGAAGCCCTGATTCGAAACCGCAGCTAGTCTTCTCTCCCTGGCCACCTGCGG 997  
Db 1057 CAGATCTGAGGGCTCTGGCAAGAAATCTCTAGTTTCTGCTGTCTTCTCCCTGGGCTTCACTG 1116  
Qy 998 CTGTCTCTCTCGCCACGGGGGCTGGGCAATGTGATCCCGCTCTACTCTGACCCGGGCCC 1057  
Db 1117 CTGTGGCTTTGTACAGGGCTCCCTGGCTCTGTGGGCTCGGCAATTCCTGCTGCTTCCC 1176  
Qy 1058 AAGTTGTGCA---GAAGACAGCAGAGACGTGCAACAGCCCGCCCTGTGGGGCCCAAGACA 1114  
Db 1177 GGTGTGCTCTGGGGAGAGCCCAACCCCTGCTTCCCGGAGACTCTCTCTCTCTCTGACA 1236  
Qy 1115 GCTCATCTTTGGGGCCATCACTGCTTTACGGGATTTCTGGGCGTGGTCAACGGGGGCGAG 1174

Db 1237 GTCTCATCTTTGGACTCATACCTGCTCCCTGACCGGAGTCTCGGTGTGGGCTTGGGCTGGGTGG 1296  
Qy 1175 GAGCCACCGCTGTGTGCGCTCTAGAGACCCAGCGGGCGAGCCCACTGCTGTGTGTCGCTGG 1234  
Db 1297 AGATCAGCGCGGCTCGGCCACTCCAAACCCCGGGCTGATCCCTGTGTCTGTGCCACTG 1356  
Qy 1235 GCATGCTGGGCTCTGCCATCTTCACTGCTGATCTTGTGGTGTGCGGAGAGAGAGATCG 1294  
Db 1357 GCCTCCTGGGCTCTGCCATCTTCTCTTCTGCTGCTGCTGCGGCGGTGTAGCATCG 1416  
Qy 1295 TAGGAGCCTATATCTGTATCTTCTGCGGGAGAGCTGCTGTTCTTAACCTGGGCGCATCA 1354  
Db 1417 TGGCCACTATATTTTCACTTTCATGTGAGAGAGACCTCTCTCATGAACTGGGCGATCG 1476  
Qy 1355 CTGCAGACATCTCATGTACGTGTGTCATCCCAAGCGCGGCGGCACTGCGCTGCGCTTGC 1414  
Db 1477 TGGCGGACATCTGCTGTACGTGTGATCCCTACCGAGGCTCCACCGCGGAGGCTTCC 1536  
Qy 1415 AGACCTTCACTCCACCTGCTGGGGAGCGCGGAGCCCTCACTCATTTGGCTTTATCT 1474  
Db 1537 AGATCGTGTCTGCCACCTGCTGGGTGATGTGTGGGAGCCCTACCTCATTTGGCTGTATCT 1596  
Qy 1475 CAGACCTGATCCCGCAGAGCACTAAGGACTCCCGCTCTGGGAGTCTCTGAGCTGGGCT 1534  
Db 1597 CTGACCGCTCGCGCGGAATGCGCCCTCTCTTCTTGTGCGAGTTCGGGCTCTGCAGT 1656  
Qy 1535 ACGCGCTCATCTCTGCGCTTTCGTGCTGTGTCCTGGGCGGATGTTCTTCTCCCGCACTG 1594  
Db 1657 TCTCGCTCATGCTCTGCGGCTTGTGTGGGCACTTGGGCGGAGAGCTTCTTCTGGGACCG 1716  
Qy 1595 CGCTCTTCTGCTCAGGACCGCGCGAGGCTGTGAGCAGACCTGGGG 1642  
Db 1717 CCATCTTCAATTAGGCGGACCGCGCGGGGACAGCTGCAGTGCAGG 1764

RESULT 8

BM147219  
LOCUS  
DEFINITION  
TCAAP1Q10621 Pediatric acute myelogenous leukemia cell (FAB M1)  
Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP1062, mRNA  
sequence.  
ACCESSION  
BM147219  
VERSION  
BM147219.1 GI:17166851  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
1 (bases 1 to 543)  
AUTHORS  
Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R., Jr.,  
Gonatne, P.H., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.  
Pediatric Leukemia cDNA Sequencing Project (2001)  
TITLE  
Unpublished (2001)  
JOURNAL  
COMMENT  
Contact: Dr. Judith F. Margolin  
Texas Children's Cancer Center and Human Genome Sequencing Center  
at Baylor College of Medicine  
1102 Bates, MC3-3320 Houston, TX 77030, USA  
Tel: 832-824-4536  
Fax: 832-825-4038  
Email: clones@ccc.org  
Seq primer: M13 primer.  
Location/Qualifiers

FEATURES

source  
1..543  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="TCAAP1062"  
/sex="male"  
/tissue\_type="leukopheresis"  
/cell\_type="myeloid cell"  
/dev\_stage="pediatric 6 years"  
/lab\_host="DH10B"  
/clone\_lib="Pediatric acute myelogenous leukemia cell (FAB

M1) Baylor-HGSC project-TCAA"  
/note=Vector: lambda pSB; Site 1: BamHI; Site 2: EcoRI;  
First strand cDNA was primed with an anchored  
XhoI-oligo(dT) primer [5'-GGAGACTCGAGCGCGGAGGAG(T)VN  
3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand  
was primed with a BamHI-dC primer  
[5'-AGAGACTCGATCCGCGCGCAATAATAAT(C) 3'].  
Double-stranded cDNA was then digested with BamHI and XhoI  
and directionally cloned into the BamHI and SalI sites of  
lambda pSB vector. Library was then through one round of  
normalization. Library was constructed by Wei Yu at RIKEN  
of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,  
Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M,  
Schneider C, Hayashizaki Y, High efficiency selection of  
full-length cDNA by improved biotinylated cap trapper.,  
DNA Res 4: 1, 61-6, Feb 28, 1997")

ORIGIN

Query Match 25.3%; Score 524.2; DB 4; Length 543;  
Best Local Similarity 98.9%; Pred. No. 4e-96;  
Matches 537; Conservative 0; Mismatches 5; Indels 1; Gaps 1;  
Qy 1037 CGCTCTACCTGCACCGCGCCCAAGTTCTGCAGAGACAGCAGACAGCTGCAACGCCCGC 1096  
Db 1 CGCTCTACCTGCACCGCGCCCAAGTTCTGCAGAGACAGCAGACAGCTGCAACGCCCGC 60  
Qy 1097 CCTGTGGGGCCAAAGACAGAGCTCATCTTTGGGGCCATCAGCTGTTACGGGATTTCTGG 1156  
Db 61 CCTGTGGGGCCAAAGACAGAGCTCATCTTTGGGGCCATCAGCTGTTACGGGATTTCTGG 120  
Qy 1157 GCGTGTTCACGGGGGCGAGGA-GCCACGCGCTGTGTGCGCCCTGAAAGACCCAGCGGGCGCAC 1215  
Db 121 GCGTGTTCACGGGGGCGAGGATTTTCACGCGCTGTGTGCGCCCTGAAAGACCCAGCGGGCGCAC 180  
Qy 1216 CCACGTGTGTGCGGTGGGCGATGCTGGGCTCTGCCATCTTTCATCTGCTGATCTTCTGTG 1275  
Db 181 CCACGTGTGTGCGGTGGGCGATGCTGGGCTCTGCCATCTTTCATCTGCTGATCTTCTGTG 240  
Qy 1276 GCTGCCAAGAGCAGCATCTGAGAGCCTATATCTGTATCTTCTGTCGGGAGACCGCTGCTG 1335  
Db 241 GCTGCCAAGAGCAGCATCTGAGAGCCTATATCTGTATCTTCTGTCGGGAGACCGCTGCTG 300  
Qy 1336 TTTTCTAACTGGGCGCATCACTGCAGACATCTCTCATGTACGTGTGTCTATCCCGCGCGCGC 1395  
Db 301 TTTTCTAACTGGGCGCATCACTGCAGACATCTCTCATGTACGTGTGTCTATCCCGCGCGCGC 360  
Qy 1396 GCACGTGCGGTGGCTTTCAGAGCTTCACTCCACCTGCTGGGGAGACCGCGGAGCGCC 1455  
Db 361 GCACGTGCGGTGGCTTTCAGAGCTTCACTCCACCTGCTGGGGAGACCGCGGAGCGCC 420  
Qy 1456 TACCTCATTTGGCTTTATCTCAGACCTGATCCGCGCAGAGCACTAAGGACTCCCGCTCTGG 1515  
Db 421 TACCTCATTTGGCTTTATCTCAGACCTGATCCGCGCAGAGCACTAAGGACTCCCGCTCTGG 480  
Qy 1516 GAGTTCCTGAGCCTGGGCTACGGGCTCATCTCTGCGCTTTTCTGCTGTGTGCTCTGGGCGGC 1575  
Db 481 GAGTTCCTGAGCCTGGGCTACGGGCTCATCTCTGCGCTTTTCTGCTGTGTGCTCTGGGCGGC 540  
Qy 1576 ATG 1578  
Db 541 ATG 543

RESULT 9

BE285645  
LOCUS  
DEFINITION  
601097128F1 NCI CGAP\_Mam5 Mus musculus cDNA clone IMAGE:3495966 5',  
mRNA sequence.  
ACCESSION  
BE285645  
VERSION  
BE285645.1 GI:9163408  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1075)  
NIH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM8547 row: 9 column: 07  
High quality sequence stop: 624.  
Location/Qualifiers  
1. .1075  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="mix FVB/N, CS7BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3495966"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH108"  
/clone\_lib="NCI CGAP Mam5"  
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"

Query Match 24.7%; Score 512.4; DB 2; Length 1075;  
Best Local Similarity 82.2%; Pred. No. 1e-93;  
Matches 661; Conservative 0; Mismatches 136; Indels 7; Gaps 5;  
QY 790 GCAGCCGAGACTGGCATTGCGGATTCGGGTTGCCCTCTCTCTGGGATGATCAGGA 849  
DB 1 GCAGCTGAGACTGGCATTGGGCGCTGCGGTTTCCCGCTCTCTGGGATGATCAGGA 60  
QY 850 ACATCTATCTCTCTGTTGCCAGCCACTAAGGGGTATCGGACCATCGGACAGCTCGGGAC 909  
DB 61 ACATCTATCTCTCTGTTGGTTCAGCCACTAAGAGAGGCCATGCTGATCACTTGGGGG 120  
QY 910 CAGCTCAAGGCCCGGACCTCATGGCTCCGAGATATGAAGGCCCTGTATTCGAAACCGCAGC 969  
DB 121 CAGCTCAAGGACGAGACTCTCTGGCTTCGAGACATGAAGGCCCTGTATTCGAAACCGCAGT 180  
QY 970 TACGTTCTCTCTCTCGGCGACATCGGCTGTCTCTTCGACCGGGGCTCTGGGATG 1029  
DB 181 TACGTTCTCTCTCTCGGCGACATCGGCTGTCTCTTCGACCGGGGCTCTGGGATG 240  
QY 1030 TGGATCCGCTCTACCTGCAACCGGCCCAAGTTGTGTCAGAGAGCAGCAGAGACTGCAAC 1089  
DB 241 TGGATCCGCTCTATCTTCACCGTGCTCAAGTTGTACAAAGACAGCAGAGACTTGAAC 300  
QY 1090 AGCCCGGCTCTGGGGGCAAGACAGCCTCATCTTTGGGGCCATCCTGCTTTACGGGA 1149  
DB 301 AGCCCGGCTCTGGAGCAAGACAGCCTCATCTTTGGAGCCATTCCTGCTTTACTGGC 360  
QY 1150 TTTCTGGGCGTGGTCAAGGGGAGAGCCAGCGGCTGGTGGCCCTGAGAGCCAGCCGG 1209  
DB 361 TTTCTGGGCGTAGTCACGGGCGGGAGCCACTCGTTGGTGGCCCTGCGGACTCAACGT 420  
QY 1210 GCCGACCCACAGTGTGGCGTGGGATGCTGGGCTCTGCACTCTTCATCTGCTGATC 1269  
DB 421 GCTGACCCATTTGGTGTGTGTGGGATGCTGGGATCCGCGCATCTTCATCTGCTCATC 480  
QY 1270 TTCGTGGTGCACAGAGCAGCATCGTAGGAGCCCTATATCTGTATCTTCGTGGGGAGACG 1329

Db 481 TTTGTGGCTGCCAAGACCAGCAATTGTAGGCGCTTATATCTGCAATTTTGTGGAGAGACA 540  
QY 1330 CTGCTGTTTTTAACCTAGCGGCATCACATGCA-GACATCTCATGTACGTGGTCAATCCCCAC 1388  
Db 541 CTGCTGTTTTTAACCTAGCGGCATCACATGCAAGCAATTTCTCATGTCTGTGGTCAATCCCCAC 600  
QY 1389 -GCCGCGCGCCACTGCGGCTGGCCCTTGACAGAGCTTCACTCCACCTGCTGGGGAGCGCG 1447  
Db 601 TCCGCGGAGCAGCTGCTGTGGCCCTGCAGAGGTT-AACTTCCATCTGCTGGGGGA--CCT 657  
QY 1448 GGAGCCCTACTCATTTGGCTTTATCT-CAGACCTGATCCGCCAGAGACACTAAGGACTCC 1506  
Db 658 GGAAGCCCTAACTCATTTGCGTTTATCTCCGACCTAATTTCCGACAGAGCCGAGATCCC 717  
QY 1507 CCGTCTGG-GAGTTCTGAGCGCTGGGCTACGCTCATGCTCTGCCCTTTCGTGCTGT 1565  
Db 718 CGTCTGGGGAGCTCCCTGAGCCTGGGTATGCCCTGTGTGTGCCCTTTGTGCGGGA 777  
QY 1566 CTTGGGCGCATGTTTCTCTCGC 1589  
Db 778 CACGCGGGCAGCATTTGTTCCC 801

RESULT 10  
BG243478 687 bp mRNA linear EST 13-FEB-2001  
LOCUS 602356937F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:448585 5',  
DEFINITION mRNA sequence.  
ACCESSION BG243478  
VERSION BG243478.1 GI:12753293  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 687)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM10327 row: 1 column: 10  
High quality sequence stop: 650.  
Location/Qualifiers  
1. .687  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:448585"  
/tissue\_type="tumor, biopsy sample"  
/dev\_stage="3 months, virgin"  
/lab\_host="DH108"  
/clone\_lib="NCI\_CGAP\_Mam1"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

FEATURES  
source  
Query Match 24.8%; Score 509.8; DB 4; Length 687;  
Best Local Similarity 89.3%; Pred. No. 3.4e-93;  
Matches 594; Conservative 0; Mismatches 67; Indels 4; Gaps 4;  
QY 960 AAACCCGAGTACGTCTTCTCCCTCCCTGGCCAGCTCGGCTGTCTCTCCACAGGGGC 1019



AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical  
College of Wisconsin  
cDNA Library Preparation: Express Genomics  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM15381 row: 0 column: 16  
High quality sequence stop: 672.  
Location/Qualifiers  
1. .815  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="IMAGE:7320138"  
/sex="both"  
/tissue\_type="Brain - Pooled from several tissues from one  
or more individuals"  
/lab\_host="DH10B Tona"  
/clone\_lib="NIH MGC 255"  
/notes="Organ: brain/CNS; Vector: pExpress-1; Site 1:  
ECORV; Site 2: NotI; RNA obtained from brain tissue of 8  
wk old animal. Tissues were snap-frozen and kept at -80C  
before RNA extraction and purification (Tri-reagent  
method). cDNA was primed using oligo-dT primer:  
5'-pGACTAGTCTAGATCGGAGCGCGCC(T)25-3' and cloned into  
the ECORV/NotI sites of pExpress-1. Size-selection >1.25kb  
resulted in an average insert size of 1.7 kb. This primary  
library is a normalized (primary library is NIH MGC 254)  
and was constructed by Express Genomics (Frederick, MD).  
Note: this is a NIH\_MGC library"

ORIGIN  
Query Match 24.3%; Score 502.6; DB 7; Length 815;  
Best Local Similarity 89.4%; Pred. No. 1e-91;  
Matches 541; Conservative 0; Mismatches 64; Indels 0; Gaps 0;  
QY 1030 TGGATCCCGCTTACTGTCACCGCCCAAGTTGTGCGAGAGACAGAGAGCTGCAAC 1089  
DB 1 TGGATCCCGCTTACTGTCACCGCTGCTCAAGTTGTACAAAGAGACAGAGAGCTTGCAC 60  
QY 1090 AGCCGCGCTGTGGGCGCAGGACAGCCTCATCTTTGGGGCCATCACCCTGTTTACGGGA 1149  
DB 61 AGCCGCGCTGTGGGCGCAGGACAGCCTCATCTTTGGGGCCATCACCCTGTTTACGGG 120  
QY 1150 TTTCTGGGCGTGTGTCAGGGGCGAGGACCGCTGTGTGCGCCCTGAAAGACCCAGCGG 1209  
DB 121 TTTCTGGGCGTGTGTCAGGCGCGGAGCCACTCGCTGTGTGCGCTTGGCGCACTCAGCG 180  
QY 1210 GCGACCCACTGTGTGTGCGGTGGGATCTGCGGCTCTGCCATCTTCATCTCGCTGATC 1269  
DB 181 GCTGACCCCTTGTGTGCGCGTGGGATCTGCGGATCTGCCATCTTCTATCTGCTCATC 240  
QY 1270 TTGCTGCTGCCAGAGCAGCATCTGTAGGAGCCTATATCTGTATCTTCGTCGGGGAGAG 1329  
DB 241 TTTGTGCTGCCAAGACAGCATCTGTGGGCGCCTATATCTGCACTCTTTGTGGGAGACA 300  
QY 1330 CTGCTGTTTCTTAACCTGGGCACTCAGTCGAGACATCTCTATGTACGTGGTGCATCCCAAC 1389  
DB 301 CTGCTGTTTCTTAACCTGGGCACTCAGTCGAGACATCTCTATGTATGTGTGTCATCCCA 360  
QY 1390 CGGCGCGCCACTGCGGTGGCTTGCAGAGCTTCACTCCACCTGCTGGGGGAGCGCGG 1449  
DB 361 CGAGAGCCACTGCTGTGGGCTTGCAGAGCTTCACTCCATCTGCTGGGGGAGCGCTGGA 420

QY 1450 AGCCCTACCTCATTTGGCTTTATCTCAGACCTGATCCGCCAGAGCACTAAGGACTCCCCG 1509  
DB 421 AGCCCTACCTCATTTGGCTTTATCTCGACCTATCCGCCAGAGCACCAGGACTCCCCA 480  
QY 1510 CTCTGGGAGTTCTCGAGCCTGGGCTACCGGCTCATGCTCTGCCCTTTCTGTCGTGCTG 1569  
DB 481 CTCTGGGAGTTCTCGAGCCTGGGCTATGCCCTCATGCTGTGCCCTTTTGTGTGCTGCT 540  
QY 1570 GCGGCACTGTTCTTCCTCGCAGCTGGCTCTTCTGTCAGGACCGCGCCAGGCTGAG 1629  
DB 541 GGTGCACTGTTCTTCCTTGGCCACTCTCTCTTCTTCAGGATCGTGCCAGGCTGAG 600  
QY 1630 CAGCA 1634  
DB 601 CAGCA 605  
RESULT 13  
BG974364  
LOCUS 602844132F1 NCI\_CGAP\_Mam4 Mus musculus cDNA clone IMAGE:4979959 5',  
DEFINITION mRNA sequence.  
ACCESSION BG974364  
VERSION BG974364.1 GI:14362001  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 724)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth  
Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10978 row: p column: 08  
High quality sequence stop: 719.  
Location/Qualifiers  
1. .724  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="NMRI"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4979959"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="5 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Mam4"  
/note="Organ: mammary; Vector: pCMV-SORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Priscilla Furth,  
NIH Reference for transgenic model: Li et al., Cell Growth  
and Differentiation 7, 3-11 (1996)."

ORIGIN  
Query Match 24.1%; Score 498.8; DB 4; Length 724;  
Best Local Similarity 85.5%; Pred. No. 5.9e-91;  
Matches 590; Conservative 0; Mismatches 97; Indels 3; Gaps 3;  
QY 691 TTCACCAAGAACACAGCGTACGCTACGCTGTC-CGTCTTCTTACTTCGCCATCCACTGGG 749  
DB 34 TTACCAAGNACACAGCATGTGCTATCTAGTCTTCTTATTTTGCCATCCCTGGG 93



QY 750 CAGTGGCTGGCTACATTACTGCTCCAGCTGAGCAGCAGCGGAGCTGGCACTG 809  
Db 94 CAGTGGCTGGCTATATACAGGTTCCAGCTGAGCAGCAGCTGGAGACTGGCACTG 153  
QY 810 GGCATTGGGGTGTCCCTCG-TCTGGGTCATGATCAGAGGAACACTCATCTCATCTGG 868  
Db 154 GGCCTTGGGGTGTCCCTCGATCTGGGTCATGATCAGAGGAACACTCATCTCATCTGG 213  
QY 869 TCCAGGCACTAAAGGGGTGATCCGACAGCTCGGGACAGCTCAAGGCCCGGACCT 928  
Db 214 TCCAGGCACTAAAGGGGTGATCCGACAGCTCGGGACAGCTCAAGGCCCGGACCT 273  
QY 929 CATGGCTCCGAGATATGAAGCCCTGATTCGAACCGAGCTAGCTCTTCTTCTCCCTGG 988  
Db 274 CTTGGCTTCGAGACATGAAGCCCTGATTCGAACCGAGCTAGCTCTTCTTCTCCCTGG 333  
QY 989 CCAGTGGCTGTCTCTTGGCCACGGGGCCCTGGGTCATGATCCCGCTCTACCTGC 1048  
Db 334 CCATCTCGCTGTCTCTTGGCCACGGGGCCCTGGGTCATGATCCCGCTCTATCTTC 393  
QY 1049 ACCGGCCCAAGTTGTGCAAGAGACAGCAGAGCTGCAACAGCCCGCTGTGGGCCA 1108  
Db 394 ACCGTCTCAAGTTGTACAAAGACAGCAGAGCTTGCACAGCCCGCTGTGGGCCA 453  
QY 1109 AGGACAGCTCATCTTTGGGCCATACCTGCTTTTACGGGATTTCTGGGGTGTGACGG 1168  
Db 454 AAGACAGCTCATCTTTGGAGCCATTACCTGCTTTTACTGGCTTCTGGGGTGTGACGG 513  
QY 1169 GGGCAGAGCCACGG-CTGGTGGCTGAGAGCCAGCGGGCGGACCCACTGGTGTGT 1227  
Db 514 GCGGGGAGCCACTGATGGTGGCTGCGGCTGCGGACTCAAGCTGTGACCCACTAGGTGT 573  
QY 1228 GCGGTGGGCTGCTGGGCTGTGCCATCTTATCTGCTGATCTTCTGGGCTGCCAAGAGC 1287  
Db 574 GCTGTGGGCTGCTGGGATCGGCATCTTATCTGCTGATCTTCTGGGCTGCCAAGAGC 633  
QY 1288 AGCATCTAGAGCTATATCTGATCTTCTGGGAGAGCTGCTGTTTTCTAACTGG 1347  
Db 634 AGCATAGTGGCGCTTATCTGATCTGAAGTGAAGGAGAGACATGCTGATTACTAACTGG 693  
QY 1348 GCATCACTGCAGACATCTCTCATGTG 1377  
Db 694 GCATCACTGCAGACATCTCTCATGTATGTG 723

RESULT 14  
AY403401 1587 bp DNA linear GSS 12-DEC-2003  
LOCUS Homo sapiens HCM1545 gene, VIRTUAL TRANSCRIPT, partial sequence,  
DEFINITION genomic survey sequence.  
ACCESSION AY403401  
VERSION AY403401.1 GI:39759384  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1587)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 1587)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
Direct Submission  
TITLE Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA  
This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source 1..1587  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..>1587  
/locus\_tag="HCM1545"  
ORIGIN  
Query Match 23.9%; Score 495.2; DB 9; Length 1587;  
Best Local Similarity 58.4%; Pred. No. 3.3e-90;  
Matches 799; Conservative 0; Mismatches 562; Indels 6; Gaps 2;  
QY 282 GGGCCCGGCGCTCAGCAGCCGCAACCGGCGGCTTGGGCGGGGGGCGGCGCGC 341  
Db 126 GGACACAGAGGGGTGCGCGGCATCACCGGCGCTGCTCCCGGCGCTCTCATAGT 185  
QY 342 CGCCATCTCAGCTTGGGCAACGCTGCTCAACTACTCGACAGGTACACCGTGGCAGCGT 401  
Db 186 GCGGTGCTGCTGTACATCAATCTCTGAATACATGACCGCTTACCGTGGTGGCGT 245  
QY 402 CTTTGTGACATCAGCAGCACTTTGGGGTCAAGACCGAGGGCCCGGCTGTGCAATC 461  
Db 246 CTTTCCGACATCGAGCAGTTCTCAACATCGGGGACAGTAGCTCTGGGCTCATCCAG 305  
QY 462 AGTGTTCATCTGAGCTTTCATGCTGGCTGCCCCCATCTTGGCTACCTGGGCGGCGCTT 521  
Db 306 CGTGTTCATCTCCAGTTACATGTTGGCACCTGTGTTGGTACCTGGTGGTACAGGA 365  
QY 522 CAACAGGAAGTGAATCTCAGCTGCGGCATTTTCTTCTGCTGGCGCGTCACTTCTCCAG 581  
Db 366 CAATCGGAATGATCTCATGTGCGGGGCAATGCTTCTGCTCCCTGTGTGACACTGGGTC 425  
QY 582 CTCCTTCATTTCCAGCAGTACTTCTGGCTGTGGTCTCTGTCTCCCGGGGCTGTGGGAT 641  
Db 426 ATCCTTCATCCCGGAGAGCATTTCTGGCTGCTCTCTGACCCCGGCGCTGTGGGGT 485  
QY 642 CGGGGAGGCGAGCTACTCCACCATCGCCCCCATCTATCTTGGGACCTCTTCCACAGAA 701  
Db 486 CGGGGAGGCGAGTTATTCCACCATCGCGCCCACTCTCATTTGCCGACCTCTTTGTGGCCGA 545  
QY 702 CACGCTGAGCTCATGCTGCTTCTTCTTCTTCTGCTTCCCTGCGGCTGAGTGGCTGGG 761  
Db 546 CAGCGGAGCGGAGTCTGAGCATCTTCTTCTTCTGCTTCCGCTGGGCTGAGTGTCTGGG 605  
QY 762 CTACATTTACTGGCTCCAGCGTGAAGCAGGCGGAGACTGGCAGCTTGGGCAATTTGGGCT 821  
Db 606 CTACATTTGAGGCTCCAAAGTGAAGGATATGGCTGGAGACTGGCAGCTGGGCTCTGAGGGT 665  
QY 822 GTCCCTGTCTGGGATGATCAAGGAACATCTCTCTCATCTTCTGTGTCCAGGCACTAA 881  
Db 666 GACACCGGGTCTAGGAGTGGTGGCGCTTCTGCTGCTTCTTCTGCTGAGTGGGAGCGGCC 725  
QY 882 AAGGGGTCTGCGGACCGAGCTCGGGGA---CCAGCTCAAGCGCGCGGACCTCATGGCTCCG 938  
Db 726 AAGGGAGCGCTGGAGCGGCACCTCAGATTTGGCCACCCCTGAACCCACCTGCTGGTGGGC 785  
QY 939 AGATATGAAGGCCCTGATTCGAAACCGCAGCTACGTCTTCTCTCTCTGCGGCACTGGC 998  
Db 786 AGATCTGAGGGCTCTGCAAGAAATCTAGTTTTCGTCCTGCTTCTTCTCTGGCTTCACTGC 845  
QY 999 TGTCTCTTTCGCAACGGGGCCCTGGGCTGTCATCTCCGCTCTTCTGCTGACCGCGGCCA 1058  
Db 846 TGTGGCTTTTGTACCGGGCTCCCTGGCTCTGTGGGCTCCCGCATTCCTGCTGGCTTCCCG 905  
QY 1059 AGTTGTGCA---GAAGACAGCAGAGCTGCACAGCCCGCCCTGTGGGCGCAGGACAG 1115  
Db 906 CGTGGTCTTGGGGAGACCCACCTGCTTCCCGGAGACTCTCTGCTTCTTCTCTGACAG 965  
QY 1116 CTTCTATTTTGGGGCCATCACCTGCTTTTACGGGATTTCTGGGCTGTGTGTCACGGGGGCGAG 1175

```
Db 966 TCTCATCTTTGGACTCATCACCTGCTGACCGAGTCTCGGTGTGGGCTGGGTGTGGA 1025
Qy 1176 AGGCAGCGCTGGTGGCTGAGACCCAGCGGGCGGACCACTGGTGTGGTGGTGG 1235
Db 1026 GATCAGCCCGGCTCGGCTCAACCCCGGGCTGATCCCTGGTGTGGTGGTGG 1085
Qy 1236 CATGCTGGGCTGCGCATCTTTCATCTGCTGATCTTCTGCTGGTGGTGGTGGTGG 1295
Db 1086 CTTCTTGGGCTTGGACCTTCTCTTCTGCTGCTGCTGGTGGTGGTGGTGGTGG 1145
Qy 1296 AGGAGCTATATCTGATCTTCTGCGGGAGAGCGTCTGTTTCTTAACCTGGGCCATCAC 1355
Db 1146 GGCACCTTATATTTTCATCTTCAATGGAGAGACCTCTCTGTCATGAATGGGCCATCGT 1205
Qy 1356 TGCAGACATCTCATGTAGTGTATCCCGACCGCGCGGCCGACCTGCGGCTGGCTGCA 1415
Db 1206 GCGCGACATCTGCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1265
Qy 1416 GAGCTTCACTCCCACTGCTGGGGACGCGCGGAGCCCTACCTCATTTGGCTTTATCTC 1475
Db 1266 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1325
Qy 1476 AGACCTGATCGCGCAGACGACTAAGGACTCCCGCTCTCGGAGTTCCTGAGCCTGGGCTA 1535
Db 1326 TGACCGCTGCGCGGAACTGGCCCCCTCTTCTTGTCCGAGTTCGCGGCTCGCAGTT 1385
Qy 1536 CGGCTCATGCTCTGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1595
Db 1386 CTGCTCATGCTCTGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1445
Qy 1596 GCTCTTCTGCTGAGCGCGCGGCTGAGCAGCAGCTGGGGG 1642
Db 1446 CATCTTCACTGAGCGCGCGGCGGCGAGCTGCGAGCTGCGAGG 1492
```

```
RESULT 15
LOCUS CN666419
DEFINITION A0840B03-5 NIA Mouse E13.5 whole embryo cDNA library (Long) Mus
ACCESSION CN666419
VERSION 1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 579)
AUTHORS Sharov,A.A., Piao,Y., Matoba,R., Dudekula,D.B., Qian,Y.,
VanBuren,V., Falco,G., Martin,P.R., Stagg,C.A., Bassey,U.C.,
Wang,Y., Carter,M.G., Hamatani,T., Alba,K., Akutsu,H., Sharova,L.,
Tanaka,T.S., Kimber,W.L., Yoshikawa,T., Jaradat,S.A., Pantano,S.,
Nagaraja,R., Boheler,K.R., Taub,D., Hodes,R.J., Longo,D.L.,
Schlessinger,D., Keller,J., Klotz,E., Kelsoe,G., Umezawa,A.,
Vescovi,A.L., Rossant,J., Kunath,T., Hogan,B.L., Curci,A.,
D'Urso,M., Kelsoe,J., Hide,W. and Ko,M.S.
Transcriptome analysis of mouse stem cells and early embryos
PLOS Biol. 1 (3), 410-419 (2003)
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Caswell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsum.grc.nia.nih.gov
Plate: A0840 row: B column: 03
Seq primer: M13 Reverse
High quality sequence stop: 579
POLY(A)=No. Location/Qualifiers
1..579
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
```

```
/db_xref="niaEST:A0840B03-5"
/db_xref="taxon:10090"
/cloned="NIA:A0840B03 IMAGE:30761006"
/tissue_type="whole embryo including extraembryonic
tissues at 13.5-days postcoitum"
/dev_stages="E13.5"
/lab_host="DH10B"
/clone_lib="NIA Mouse E13.5 whole embryo cDNA library
(Long)"
/notes="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
Site_2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsum.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
RNAs were extracted from 1 embryo at 13.5-days postcoitum.
Double-stranded cDNAs were synthesized with an Oligo(dT)
primer [Invitrogen]:
5'-pGACTAGTTCTAGATCGGAGCGCGCCCTTTTTTTTTTTT-3' from
3ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lona-linker LL-SalI, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer SalI-4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 3.0Kb. The library was
constructed by Yulan Piao."
```

ORIGIN

```
Query Match 23.3%; Score 483; DB 7; Length 579;
Best Local Similarity 89.6%; Pred. No. 9.6e-88;
Matches 519; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 1049 ACCGCGCCCAAGTCTGCAGAGAGCAGCAGAGAGCTGCAACAGCCGCCCTCTGGGGCCA 1108
Db 1 ACCGTGCTCAAGTTGTACAAAGAGCAGCAGAGACTTGCACAGCCGCCCTCTGGAGCCA 60
QY 1109 AGGACAGCCTCATCTTTGGGGCCATCACCCTGCTTTACGGGATTTCTGGGCGTGGTCA 1168
Db 61 AAGACAGCCTCATCTTTGGAGCCATTACCTGCTTTACTGGCTTCTGGGCGTAGTCACGG 120
QY 1169 GGGCAGGAGCCACCGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1228
Db 121 GCGCGGAGGCCACTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180
QY 1229 CCGTGGGCGATGCTGGGCTGCGCATCTTCATCTGCTGATCTTCGTGGCTGCAAGAGCA 1288
Db 181 CTGTGGGCGATGCTGGGATCCCGCATCTTCATCTGCTGCTTCATCTTTGGGCTGCAAGAGCA 240
QY 1289 GCATCGTAGGAGCCATATCTGTATCTTCGTGGGAGACGCTGCTGTTTCTTAACCTGGG 1348
Db 241 GCATTTAGGCGCTTATATCTGCACTTTTGTGGAGAGACACTGCTGTTTCTTAACCTGGG 300
QY 1349 CCATCACTGCGAGACATCTCATGTACGTGGTGCATCCACAGCGGGCGGCCACTGCGGTGG 1408
Db 301 CCATCACTGCGAGACATCTCATGTATGTGTGCATCCCACTCGGGAGGAGCACTGCTGTGG 360
QY 1409 CTTTCAGAGCTTACCTCCACCTGCTGGGGAGCGCGGGAGCCCTACCTCATTTGGCT 1468
Db 361 CCCTCGAGAGCTTACCTCCACCTGCTGGGGAGCGCTGGAAGGCCCTTACCTCATTTGGCT 420
QY 1469 TTATCTCAGACCTGATCGCCAGAGCACTAAGGACTCCCGCTCTGGGAGTTCTTGAGCC 1528
Db 421 TTATCTCGGAGCTAATCGCCAGAGCAACAGGACTCCCGCTCTGGGAGTTCTTGAGCC 480
QY 1529 TGGGCTACGCGCTCATGCTCTGCCCTTCTGCTGGTGTGGTGGGCGGATGTTCTTCTCTG 1588
Db 481 TGGGCTATGCGCTCATGCTGTGCCCTTTTGTGCTGCTGCTGGTGGGCGATGTTCTTCTCTG 540
```

Qy 1589 CCACTGGCGCTCTTCTTCGTACGCGACCGCGCCAGGGCTG 1627  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 541 CCACTGGCTCTTCTTCTTCCTCAGCGACCGTGCCAAAGGCTG 579

Search completed: October 23, 2004, 12:50:41  
Job time : 6572 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 08:34:23 ; Search time 89 Seconds  
(without alignments)  
2281.357 Million cell updates/sec

Title: US-10-085-198-48

Perfect score: 2920

Sequence: 1 MMLECSAAGGAEED.....VHQGPQGTALAHRVGAS 566

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 200273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2920	100.0	566	5 ADH48764	Adh48764 NOV19 pro
2	1978	67.7	398	4 AAE04896	Aae04896 Human tra
3	1604	54.9	312	7 ADE08326	Ade08326 Novel pro
4	1298	44.5	528	4 AAM38697	Aam38697 Human pol
5	1298	44.5	528	5 ABP52155	Abp52155 Human 464
6	1298	44.5	528	7 ADD37478	Add37478 Human tra
7	1298	44.5	528	8 ADL27957	Adl27957 Human 464
8	1298	44.5	528	8 ADL61327	Adl61327 Human pro
9	1256.5	43.0	507	8 ADO57285	Ado57285 Kidney de
10	1250	42.8	581	4 AAM40483	Aam40483 Human pol
11	1222	41.8	455	5 ABB90265	Abb90265 Human pol
12	1176	40.3	244	3 AAB23610	Aab23610 Human sec
13	1167	40.0	476	4 AAE06616	Aae06616 Human pro
14	1167	40.0	476	4 AAE04899	Aae04899 Human tra
15	1115	38.2	420	7 ADL08448	Adl08448 Human can
16	1088	37.3	454	4 ABUS2950	Abu52950 Human tra
17	1084	37.1	512	5 AAE16786	Aae16786 Human tra
18	1083	37.1	512	7 ADE08011	Ade08011 Novel pro
19	1054	36.1	572	5 ABB08218	Abb08218 Human mem
20	1042	35.7	221	4 AAM41959	Aam41959 Human pol
21	1030	35.3	216	4 AAM40173	Aam40173 Human pol
22	964.5	33.0	605	4 AAB67671	Aab67671 Amino aci
23	930.5	31.9	605	4 AAB67672	Aab67672 Amino aci
24	855	29.3	531	7 ADD37504	Add37504 C. elegans
25	855	29.3	531	8 ADI27938	Adi27938 Human IC4

26	779	26.7	299	2 AAW74805	Aaw74805 Human sec
27	779	26.7	299	5 ABG95256	Abg95256 Human nov
28	779	26.7	299	6 ABO34450	Ab034450 Region of
29	779	26.7	299	7 ADI23111	Adi23111 Novel hum
30	779	26.7	299	8 ADH74113	Adh74113 Human sec
31	744.5	25.5	552	4 ABB62015	Abb62015 Drosophila
32	732	25.1	144	5 ABB53205	Abb53205 Human ORF
33	672	23.0	424	2 AAW74956	Aaw74956 Human sec
34	672	23.0	424	5 ABG95415	Abg95415 Human nov
35	672	23.0	424	6 ABO34609	Ab034609 Region of
36	672	23.0	424	7 ADI23270	Adi23270 Novel hum
37	672	23.0	424	8 ADH74272	Adh74272 Human sec
38	644	22.1	229	7 ADM04605	Adm04605 Human pro
39	559	19.1	343	6 ADA11867	Ada11867 Human nov
40	553	18.9	304	6 ADA11625	Ada11625 Human nov
41	534	18.3	106	5 ABB53122	Abb53122 Human ORF
42	533	18.3	248	7 ADD78251	Add78251 Human CGD
43	531	18.2	212	3 AAB41388	Aab41388 Human ORF
44	465	15.9	256	6 ABP75842	Abp75842 Human sec
45	459.5	15.7	204	4 ABUS2951	Abu52951 Human tra

ALIGNMENTS

RESULT 1

ADH48764  
ID ADH48764 standard; protein; 566 AA.

AC ADH48764;

XX  
XX  
DT 25-MAR-2004 (first entry)

XX  
XX  
DE NOV19 protein sequence, SEQ ID 48.

XX  
XX  
KW Human; NOVX; atherosclerosis; hypertension; obesity; cancer; cytostatic;  
KW hypotensive; antiarteriosclerotic; anorectic; gene therapy; NOV19;  
KW spinster-like protein; chromosome 17.

OS Homo sapiens.

XX  
XX  
FN WO200268652-A2.

XX  
XX  
PD 06-SEP-2002.

XX  
XX  
PF 26-FEB-2002; 2002WO-US005910.

XX  
XX  
PR 26-FEB-2001; 2001US-0271646P.

PR 27-FEB-2001; 2001US-0271840P.

PR 28-FEB-2001; 2001US-0272404P.

PR 28-FEB-2001; 2001US-0272405P.

PR 28-FEB-2001; 2001US-0272410P.

PR 28-FEB-2001; 2001US-0272414P.

PR 02-MAR-2001; 2001US-0272787P.

PR 02-MAR-2001; 2001US-0272922P.

PR 02-MAR-2001; 2001US-0273048P.

PR 16-MAR-2001; 2001US-0273300P.

PR 20-MAR-2001; 2001US-0273640P.

PR 20-MAR-2001; 2001US-0278660P.

PR 30-MAR-2001; 2001US-0280039P.

PR 30-MAR-2001; 2001US-0280234P.

PR 02-APR-2001; 2001US-0280818P.

PR 12-APR-2001; 2001US-0283443P.

PR 23-APR-2001; 2001US-0285754P.

PR 24-APR-2001; 2001US-0286096P.

PR 03-MAY-2001; 2001US-0288353P.

PR 17-MAY-2001; 2001US-0291703P.

PR 31-MAY-2001; 2001US-0294834P.

PR 20-JUN-2001; 2001US-0299895P.

PR 21-JUN-2001; 2001US-0299845P.

PR 05-JUL-2001; 2001US-0303242P.

PR 13-AUG-2001; 2001US-0311981P.



stroke, cerebral neoplasms, Pick's disease, Huntington's disease and Parkinson's disease, demyelinating diseases, mental disorders including mood, anxiety, schizophrenia and seasonal affective disorder, muscle disorder including cardiomyopathy, myocarditis, polymyositis, dermatomyositis, arrhythmias and asthma and immunological disorders including AIDS, adult respiratory distress syndrome (ARDS), allergies, anemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's syndrome, systemic lupus erythematosus and other diseases including sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, Grave's disease, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers psoriasis and viral, bacterial, fungal, helminthic and protozoal infections. TRICH DNA is useful in gene therapy and in diagnostic purposes

XX SQ Sequence 398 AA;

Query Match 67.7%; Score 1978; DB 4; Length 398;  
Best Local Similarity 99.0%; Pred. No. 5.3e-170;  
Matches 384; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 152 MVAAPIFGYLDRNRKRVILSCGIFFWSAVTFSSFIPOQYFLLVLSRGLVIGIEASYS 211  
Db 1 MVAAPIFGYLDRNRKRVILSCGIFFWSAVTFSSFIPOQYFLLVLSRGLVIGIEASYS 60  
Qy 212 TIAPTIIGDLFTKNTKRLMLSVFYPAIPLGSLGYITGSSVVKQAAGDHWALRVSPVLGM 271  
Db 61 TIAPTIIGDLFTKNTKRLMLSVFYPAIPLGSLGYITGSSVVKQAAGDHWALRVSPVLGM 120  
Qy 272 ITGTLLIILVPATKRGHADQDGLKARTSWLRDMKALIRNSYVFSLSATSAVSPATGA 331  
Db 121 ITGTLLIILVPATKRGHADQDGLKARTSWLRDMKALIRNSYVFSLSATSAVSPATGA 180  
Qy 332 LGMWIPYLHRAQVQKTAETCNPPCGAKDSLIFGAIITCFTPLGVVVTGAGATRWCRK 391  
Db 181 LGMWIPYLHRAQVQKTAETCNPPCGAKDSLIFGAIITCFTPLGVVVTGAGATRWCRK 240  
Qy 392 TORADPLVCAVGMILGSAIFICLIFVAAKSSIVGAYICIFVGETLLFNSWAIADILMYV 451  
Db 241 TORADPLVCAVGMILGSAIFICLIFVAAKSSIVGAYICIFVGETLLFNSWAIADILMYV 300  
Qy 452 IPTRRATAVALQSTSHLLGDAGSPYLIIGFISDLIROSTKDSPLWEFLSLGYALMLCPFV 511  
Db 301 IPTRRATAVALQSTSHLLGDAGSPYLIIGFISDLIROSTKDSPLWEFLSLGYALMLCPFV 360  
Qy 512 VVLGGMEFLATLFFVSDRAAEQHLGE 539  
Db 361 VVLGGMEFLATLFFVSDRAAEQVQVNO 388

RESULT 3  
ADE08326  
ID ADE08326 standard; protein; 312 AA.

XX AC ADE08326;

XX DT 29-JAN-2004 (first entry)

XX DE Novel protein (useful for identifying genetic disorders) #481.

XX KW novel gene; novel protein; tissue marker; molecular weight marker;  
XX KW chromosome marker; genetic disorder.

XX OS Unidentified.

XX PN WQ0203054152-A2.

XX PD 03-JUL-2003.

XX PF 10-DEC-2002; 2002WO-US039555.

XX PR 10-DEC-2001; 2001US-0339739P.

XX PR 11-DEC-2001; 2001US-0339453P.

PR 14-MAR-2002; 2002US-0365091P.  
PR 14-MAR-2002; 2002US-0365384P.  
PR 12-APR-2002; 2002US-0372381P.  
PR 12-APR-2002; 2002US-0372615P.  
PR 22-APR-2002; 2002US-00128558.  
PR 24-APR-2002; 2002US-0376045P.  
XX (HYSE-) HYSEQ INC.  
XX PA Tang YT, Aeundi V, Goodrich RW, Ren P, Zhang J, Zhao QA, Wang J;  
XX PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;  
XX PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;  
XX WPI; 2003-569235/53.  
DR N-PSDB; ADE07415.  
XX PT New polynucleotides, useful for expressing recombinant proteins for  
XX PT analysis, characterization or therapeutic use, or as markers for tissues  
XX PT in which the corresponding protein is preferentially expressed.  
XX PS Claim 20; SEQ ID NO 1392; 1177pp; English.

XX CC The invention comprises the amino acid and coding sequences of novel  
XX CC proteins. The DNA and protein sequences of the invention are useful as:  
XX CC markers for tissues in which the corresponding protein is preferentially  
XX CC expressed; as molecular weight markers on gels; as chromosome markers or  
XX CC tags; to identify chromosomes or to map related gene positions; and to  
XX CC compare with endogenous DNA sequences in patients to identify potential  
XX CC genetic disorders. The present amino acid sequence represents a protein  
XX CC of the invention.

XX SQ Sequence 312 AA;

Query Match 54.9%; Score 1604; DB 7; Length 312;  
Best Local Similarity 99.7%; Pred. No. 2.6e-136;  
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MMCLCASAAGGAEEBEEADAERRRRRRGAQRGAGGCCGCCGARGAGGAGVSAAGDEVQTL 60  
Db 1 MMCLCASAAGGAEEBEEADAERRRRRRGAQRGAGGCCGCCGARGAGGAGVSAAGDEVQTL 60  
Qy 61 SGSVRRAPTPGTPGTGCAATAGKGAQPKASLGRGRGAAATLSLGNVLYLDY 120  
Db 61 SGSVRRAPTPGTPGTGCAATAGKGAQPKASLGRGRGAAATLSLGNVLYLDY 120  
Qy 121 TVAGVLLDIQQHFGVKDRGAGLLQSVFICSPFWAAPIFGYLGRFNRKVLSCGIFFWA 180  
Db 121 TVAGVLLDIQQHFGVKDRGAGLLQSVFICSPFWAAPIFGYLGRFNRKVLSCGIFFWA 180  
Qy 181 VTFSSSFIPQOYFWLLVLSRGLVIGIEASYSTIAPTIIIGDLFTKNTKRLMLSVFYPAIPL 240  
Db 181 VTFSSSFIPQOYFWLLVLSRGLVIGIEASYSTIAPTIIIGDLFTKNTKRLMLSVFYPAIPL 240  
Qy 241 GSGLGIVTSSVKQAAGDHWALRVSPVLGMITGTLLILVLPATKRGHADQDGLKART 300  
Db 241 GSGLGIVTSSVKQAAGDHWALRVSPVLGMITGTLLILVLPATKRGHADQDGLKART 300  
Qy 301 SWLRDMKALIR 311  
Db 301 SWLRDMKALIR 311

RESULT 4  
AAM38697  
ID AAM38697 standard; protein; 528 AA.

XX AC AAM38697;

XX DT 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 1842.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;



KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
XX WO200153312-A1.  
XX  
XX 26-JUL-2001.  
XX  
XX 26-DEC-2000; 2000WO-US034263.  
XX  
XX 23-DEC-1999; 99US-00471275.  
PR 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-00552317.  
PR 20-JUN-2000; 2000US-00598042.  
PR 19-JUL-2000; 2000US-00620312.  
PR 03-AUG-2000; 2000US-00653450.  
PR 14-SEP-2000; 2000US-00662191.  
PR 19-OCT-2000; 2000US-00693036.  
PR 29-NOV-2000; 2000US-00727344.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang JT, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
PI Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI; 2001-442253/47.  
DR N-PSDB; AAI57853.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders such  
PT as central nervous system injuries.  
XX  
XX Example 3; SEQ ID NO 1842; 10078pp; English.  
XX  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification  
XX  
XX Sequence 528 AA;  
SQ  
Query Match 44.5%; Score 1298; DB 4; Length 528;  
Best Local Similarity 53.4%; Pred. No. 2.4e-108;  
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;  
QY 53 AGDEVTQLSGSVRRAPTGP-PGTGTPGCAATAGPGCAQPKPASLGR-----GRGA-A 104  
DB 2 AGSDTAPFLSQADDPPDPVPGTGLPGSTGNPKSEPEVPDQEGQLRTGLSPGRSALI 61  
QY 105 AAISLGNVLNLYDRYTVAGVLLDIQQHFGVKDRGAGLLQSFTCSFWMVAAPFGYLGDR 164  
DB 62 VAVLCYNLLNMYDRFTVAGVLPDIEQFENIGDSSGLIQTVFTSSYMWLAPVFGYLGDR 121  
QY 165 PNRKVLSCGIFFFNSATFSSTFPQOYFWLLVLSRGVGTGEASYSTIAPTIGDLPFK 224  
DB 122 YNRKYLCCGGAFWFSFTLGSFIPGBHFVLLLTGLVGVGEASYSTIAPTIALDFA 181  
QY 225 NTRTLMLSVFYFAIPGLSGGLGYITGSSVYKQAAGDWHWALRVSPVLGMITGTLLILVPAT 284

Db 182 DQSRMLSIYFAIPVSGSLGIAGSKVQDWDHWRVTPGLGVVAVLLFLVVRP 241  
QY 285 KRHADQIGDQOLKAR-TSWLRDMKALIRNRSYVSSSLATSATGALGMWIPLYLHRA 343  
Db 242 PRGAVERHSDLPPLNPTSMWADLRALRNPSFVLSISGFTAVAFVTSGLAWAPFLURS 301  
QY 344 QVQKTAETC-NSPPCGAKDSLIFGAICTCFGLGVVTGAGATRCWLKTKTORADPLVCV 402  
Db 302 RVILGETPPCLPGDSCSSDSLIFGLITCLTGLVGVGLGVEISRRLHNSPRADPLVCAT 361  
QY 403 GMLGSAIFICLIFVAAKSSIYVAYICIFVGTLLFSNWAITADILMYVIITRATVAL 462  
Db 362 GLGSAFPLFLSLACARSIYATVIFIFIGETLLSMWAIIVADILLYYVIVITRSTABAF 421  
QY 463 QSFTSHLLGDAGSPYLIGFISDLIRQSTKDSPLMEFLSLGVYALMLCPVVLVGMFFLAT 522  
Db 422 QIVLSHLLGDAGSPYLIGLISDLIRRNPPSFLSEFRALQFSLMLCAVFGALGGAFLGT 481  
QY 523 ALFFVSDRARAEOHLGE--RRAGVR-----VVHQRG 551  
Db 482 AIFTEADRRRAQLHVQGLLHEAGSTDDRIIVVPQRG 516  
RESULT 5  
ABP52155  
ID ABP52155 standard; protein; 528 AA.  
XX  
XX ABP52155;  
XX  
XX 14-OCT-2002 (first entry)  
XX Human 46455 transporter protein SEQ ID NO:5.  
XX Human; transporter protein; anorectic; antidiabetic; anti-Parkinsonian;  
KW nootropic; neuroprotective; hypotensive; antidepressant; neuroleptic;  
KW cardiovascular; immunosuppressive; gene therapy; sugar homeostasis;  
KW obesity; diabetes; anorexia; central nervous system disorder; depression;  
KW CNS disorder; Parkinson's disease; Alzheimer's disease; hypertension;  
KW autonomic function disorder; schizophrenia; learning disorder; amnesia;  
KW memory disorder; age-related disorder; cardiovascular disorder;  
KW ischaemia reperfusion injury; restenosis; hormonal disorder;  
KW hypothyroidism; hyperthyroidism; immune disorder;  
KW chronic mucocutaneous candidiasis.  
XX  
XX Homo sapiens.  
OS WO200255701-A2.  
XX  
XX 18-JUL-2002.  
XX  
XX 17-DEC-2001; 2001WO-US049060.  
XX  
XX 15-DEC-2000; 2000US-0256240P.  
PR 18-DEC-2000; 2000US-0256588P.  
PR 21-DEC-2000; 2000US-0258028P.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
PA  
XX Curtis RAJ;  
XX  
XX WPI; 2002-590672/63.  
DR N-PSDB; ABQ74262, ABQ74263.  
XX  
XX New 67076, 67102, 44181, 67084FL and 67084alt nucleic acids and proteins,  
PT useful in e.g. treating disorders characterized by insufficient or  
PT excessive production of the polypeptides, e.g. cardiovascular or immune  
PT disorders.  
XX  
XX Claim 13; Fig 8; 364pp; English.  
XX  
XX The present sequence represents the human 46455 protein, which is a  
CC member of the transporter family. The transporter family proteins (I) of

CC the present invention have anorectic, antidiabetic, anti-Parkinsonian,  
 CC nootropic, neuroprotective, hypotensive, antidepressant, neuroleptic,  
 CC cardiovascular and immunosuppressive activities, and can be used in gene  
 CC therapy. The human transporter proteins from the present invention are  
 CC designated 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL and  
 CC 67084alt. The polynucleotides encoding (I) can be used in screening  
 CC assays (e.g. chromosome mapping, tissue typing, or in forensic biology),  
 CC predictive medicine (e.g. diagnostic assays, prognostic assays,  
 CC monitoring clinical trials, or pharmacogenetics), as surrogate markers,  
 CC and in methods of treatment (e.g. therapeutic or prophylactic). (I) are  
 CC useful for treating disorders characterised by insufficient or excessive  
 CC production of 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL or  
 CC 67084alt substrates or production of transport 8099, 46455, 54414, 53763,  
 CC 67076, 67102, 44181, 67084FL or 67084alt inhibitors. Disorders associated  
 CC with 8099 and 46455 include disorders associated with sugar homeostasis  
 CC such as obesity, diabetes or anorexia. Disorders associated with 54414,  
 CC 53763, 67076, 67102, 44181, 67084FL or 67084alt are e.g. CNS disorders  
 CC (e.g. Parkinson's disease or Alzheimer's disease), autonomic function  
 CC disorders (e.g. hypertension, depression or schizophrenia), or learning  
 CC or memory disorders (e.g. amnesia or age-related disorders),  
 CC cardiovascular disorders (e.g. ischaemia reperfusion injury or  
 CC restenosis), hormonal disorders (e.g. hypothyroidism or hyperthyroidism),  
 CC or immune disorders (e.g. chronic mucocutaneous candidiasis)  
 CC  
 CC Sequence 528 AA;

Query Match 44.5%; Score 1298; DB 5; Length 528;

Best Local Similarity 53.4%; Pred. No. 2.4e-108;

Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;

QY 53 AGDEVQTLSSVRRAPGTP-PGTGTPGCAATAGKGAQPKPASLGR-----GRGA-A 104  
 DB 2 AGSDTAPFLSQADDDPGVPTGPGTGNPKSEPEVDQGLQRTGLSPGRSALI 61  
 QY 105 AAILSLGNVLYDRYTVAGVLLDIQQHFGVDEGAGLLQSVFCSFWAAPFGYLGDR 164  
 DB 62 VAVLCYINLNYMDRFTVAGVLPDIEQFFNIGDSSGLIQTFVSSVNLAPVFGYLGDR 121  
 QY 165 FNKRVLSCGIFFWSAVTFSSSFPQQYFWLLVLSRGLVGIAGEASYSTIAPTIGDLFTK 224  
 DB 122 YNRKYLMSGGIAFWSLVTLGSSFPGBHFVWLLLTTRGLVGVGEASYSTIAPTILADLFA 181  
 QY 225 NRTMLSVFYFAIPGLSGLYGTGSKVQAAGDHWALRVSPVLGMITGTILILVLPAT 284  
 DB 182 DQSRMLSIIFYFAIPVSGGLGYTAGSKVKOMAGDHWALRVTPGLGVVAVLLFLVREP 241  
 QY 285 KRGHADQLGDLKAR-TSWLRDMKALIRNSVFPSSLATSAVSPATGALGWIPLYLHRA 343  
 DB 242 PRGAVERHSDLPPLNPTNSWADLRALARNPSFVLSLGLFTAVAVTGLSALWAPFLRS 301  
 QY 344 QVQKTAETC-NSPPCGAKDSLIFGALTCTFTGLGVVTVGAGATWCRKLTQRADPLVCV 402  
 DB 302 RVVLGETPCLPGDSCSSSLIFGLTCTGLVGLGVGEISRLRHSNPRADPLVCAT 361  
 QY 403 GMLGSAIFCLIFVAAKSSIVGAYICIFVGETLLFSNWAITADILMYVVIPTRRATAVAL 462  
 DB 362 GLLGSAFFLSLACARGSVATYIFIFIGETLLSNMNAIVADILLVYVVIPTRRSTAEAF 421  
 QY 463 QSTSHLLGAGSPYLGFTSDLIQSTKDSPLWFLSLGYALMCLPFVVVLGGMFLAT 522  
 DB 422 QIVLSHLLGAGSPYLGFTSDLIQSTKDSPLWFLSLGYALMCLPFVVVLGGMFLAT 522  
 QY 523 ALFFVSDRAEAOHLGE--RRAGVR----VVHORG 551  
 DB 482 AIFEADRRRAQLHVQGLLHEAGSTDDRIVVPQRG 516

RESULT 6

ADD37478

ID ADD37478 standard; protein; 528 AA.

XX AC ADD37478;

XX AC ADD37478;

XX AC ADD37478;

DT 15-JAN-2004 (first entry)  
 XX Human transporter 46455.  
 DE  
 XX Human; transporter; cytostatic; anorectic; antidiabetic; anticonvulsant;  
 KW gene therapy; PGC-1 associated disorder; liver tumour; obesity; epilepsy;  
 KW diabetes.  
 XX Homo sapiens.  
 OS  
 XX US2003143675-A1.  
 PN  
 XX 31-JUL-2003.  
 PD  
 XX  
 PF  
 XX 22-MAY-2002; 2002US-00154419.  
 PR 12-MAY-2000; 2000US-0204211P.  
 PR 29-JUN-2000; 2000US-0215376P.  
 PR 31-JUL-2000; 2000US-0221769P.  
 PR 19-SEP-2000; 2000US-0233790P.  
 PR 25-SEP-2000; 2000US-0235107P.  
 PR 05-OCT-2000; 2000US-0238336P.  
 PR 14-NOV-2000; 2000US-0248364P.  
 PR 15-NOV-2000; 2000US-0248878P.  
 PR 15-DEC-2000; 2000US-0256240P.  
 PR 18-DEC-2000; 2000US-0256588P.  
 PR 21-DEC-2000; 2000US-0258028P.  
 PR 22-JAN-2001; 2001US-0263169P.  
 PR 14-MAY-2001; 2001US-00858194.  
 PR 29-JUN-2001; 2001US-00895811.  
 PR 31-JUL-2001; 2001US-00919781.  
 PR 19-SEP-2001; 2001US-00957664.  
 PR 25-SEP-2001; 2001US-00964295.  
 PR 05-OCT-2001; 2001US-00972724.  
 PR 14-NOV-2001; 2001US-00002769.  
 PR 17-DEC-2001; 2001US-00024623.  
 PR 22-JAN-2002; 2002US-00055025.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA  
 XX  
 XX  
 PI Curtis RAJ, Glucksmann WA, Meyers RE;  
 XX  
 DR WPI; 2003-851783/79.  
 DR N-PSDB; ADD37477, ADD37479.  
 XX  
 PT New isolated nucleic acid, useful for preparing a composition for  
 PT treating PGC-1 associated disorders e.g. liver tumors, obesity, epilepsy  
 PT or diabetes.  
 XX  
 PS Claim 11; SEQ ID NO 55; 663bp; English.  
 CC The invention relates to an isolated nucleic acid comprising a cDNA  
 CC encoding a human transporter protein, or its complement, a sequence that  
 CC is 60 % identical to the cDNA, a fragment comprising at least 30  
 CC nucleotides of the cDNA, or a sequence encoding a fragment of the  
 CC polypeptide comprising at least 10 contiguous amino acid residues of the  
 CC cDNA. Also included are a vector comprising the novel nucleic acid  
 CC molecule, producing the polypeptide, the isolated transporter  
 CC polypeptide, an isolated antibody that specifically binds to the  
 CC polypeptide, detecting the presence of the polypeptide or nucleic acid in  
 CC a sample, a kit, identifying a compound that binds to, or that modulates  
 CC the activity of, the polypeptide, and modulating the activity of the  
 CC polypeptide. The nucleic acid is useful for preparing a composition for  
 CC treating PGC-1 (not defined) associated disorders e.g. liver tumors,  
 CC obesity, epilepsy or diabetes. The present sequence represents a novel  
 CC human transporter protein.  
 XX  
 SQ Sequence 528 AA;

Query Match 44.5%; Score 1298; DB 7; Length 528;

Best Local Similarity 53.4%; Pred. No. 2.4e-108;

Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;

```
QY 53 AGDEVQTLSGSVRRAPTGP-PGTGTPGCAATAKPGGAQPKPASIGR-----GRGA-A 104
Db 2 AGSDTAPFLSQADDDPDGVPFGTGLPGSTGNPKSEEPVQDGLQRTIGLSPGRSALI 61
QY 105 AAILSLGNVNLDRYTVAGVLLDIOQHFGVDRGAGLQSVFICSFMAAPFVGLGDR 164
Db 62 VAVLCYINLLNMDRFTVAGVLPDIEQFNIGDSSGLIQTVFISYMWLAFVFGVGLDR 121
QY 165 FNRKVLSCGIPFWSAVTFSSSFIPOQYFWLLVLSRGLVIGGEASYSTIAPTIGDLFTK 224
Db 122 YNRKYLKMGCGIAFWLSVLTGSSFIPEHFWLLLTGLVGVGEASYSTIAPTIALDFVA 181
QY 225 NTRTLMVSYFAIPGLSGGLVITGSSVKQAAGDHWALRVSPVLGMITGTLILVPAT 284
Db 182 DQSRMLSFYFAIPVGSGLGYIAGSKVDMAGDHWALRVTPGLGVWAVLLFLVVRP 241
QY 285 KEGHADQLDQLKAR-TSWLRDMKALIRNRSVVFSSLATSAVSFATGALGMWIPLYLHRA 343
Db 242 PRGAVERHSDLPPLNPTSWADRLARNPFSVLSLGFATAVFTGSLALWAPAFLLRS 301
QY 344 QVVQKTAETC-NSPPCGAKDSLIFGAITCFTGFLGVVTVGAGATRCRLKQTRADPLVCAV 402
Db 302 RVVLGETPPLCPGDCSSSDSLIFGLITCLTGLVGLGVGEISRLRHSNPRADPLVCA 361
QY 403 GMLGSNIFCLIFVAAKSSIVGAYICIFVGETLLPSNWAITADILMYVVIPTERTAVAL 462
Db 362 GLGSGAPFLSLACARGSIYAFIFIGETLLSNMNAIVADILLYVVIPTRRSTAEAF 421
QY 463 QSFTSHLLGDAGSPYLIGISDLIRQTSKDSPLWBFSLGVALMLCPFFVVLGSGMFFLAT 522
Db 422 QIVLSHLLGDAGSPYLIGISDLRNRNWPFSFLSEPRALQPSLMCAFVGLGGAFLGT 481
QY 523 ALFFVSDRARAOHLGSE--RRAGVR-----VHQRG 551
Db 482 AFIPEARRRRAQLHVQGLLHEAGSTDRTVWPQRG 516

RESULT 7
ID ADI27957
XX ADI27957 standard; protein; 528 AA.
AC ADI27957;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human 46455 protein.
XX
KW Human; 46455; ion channel family; ICF; cancer; leukaemia;
KW Alzheimer's disease; Parkinson's disease; multiple sclerosis; epilepsy;
KW hepatic disorder; cardiovascular disorder; cytostatic; neuroprotective;
KW nontropic; antiparkinsonian; hepatotropic; cardiovascular.
XX
OS Homo sapiens.
XX
PN US2003165891-A1.
XX
PD 04-SEP-2003.
XX
PF 15-MAY-2002; 2002US-00146733.
XX
PR 29-FEB-2000; 2000US-00515520.
PR 29-FEB-2000; 2000US-0185338P.
PR 03-MAR-2000; 2000US-00518866.
PR 07-APR-2000; 2000US-0195734P.
PR 11-APR-2000; 2000US-0195993P.
PR 26-APR-2000; 2000US-0199799P.
PR 19-SEP-2000; 2000US-0233537P.
PR 25-SEP-2000; 2000US-0235018P.
PR 25-SEP-2000; 2000US-0235059P.
PR 18-DEC-2000; 2000US-0256240P.
PR 21-DEC-2000; 2000US-0256588P.
PR 28-FEB-2001; 2001US-00796720.
```

```
PR 06-APR-2001; 2001US-00828035.
PR 11-APR-2001; 2001US-00833081.
PR 25-APR-2001; 2001US-00843128.
PR 19-SEP-2001; 2001US-00957683.
PR 25-SEP-2001; 2001US-00964252.
PR 25-SEP-2001; 2001US-00964256.
PR 17-DEC-2001; 2001US-00024623.
XX (MILL-) MILLENNIUM PHARM INC.
PA Curtis RAJ, Glucksmann MA, Silos-Santiago I;
PI
XX
XX WPI: 2004-069000/07.
XX N-PSDB; ADI27956, ADI27958.
XX
PT TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615, HNMDA-1, TWIK-9,
PT alpha2delta-4, 54414, or 53763 nucleic acids and proteins, useful for
PT preventing, diagnosing and treating e.g. cancers, Alzheimer's disease and
PT cardiovascular disorders.
XX
PS Example 1; SEQ ID NO 50; 638pp; English.
XX
CC The invention relates to TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611,
CC IC47615, HNMDA-1, TWIK-9, alpha-2delta-4, 54414 and 53763 ion channel
CC family (ICF) nucleic acids and proteins. The TWIK-6, TWIK-7, IC23927,
CC TWIK-8, IC47611, IC47615, HNMDA-1, TWIK-9, alpha-2delta-4, 54414 and
CC 53763 ICF nucleic acids and proteins may be used for preventing,
CC diagnosing and treating ICF-related diseases. The sequences may be used
CC to treat disorders associated with decreased expression by rectifying
CC mutations or deletions in a patient's genome that affect the activity of
CC ICF proteins by expressing inactive proteins or to supplement the
CC patients own production of ICF proteins. The proteins may also be used as
CC antigens in the production of antibodies against ICF proteins and in
CC assays to identify modulators of ICF protein expression and activity. The
CC anti-ICF protein antibodies, agonists and antagonists may be used to
CC regulate ICF protein expression and activity. The antibodies may also be
CC used as diagnostic agents for detecting the presence of ICF proteins in
CC samples (e.g. by immunoassay). The nucleic acids and proteins may be used
CC to prevent, diagnose and treat a wide variety of disorders, e.g. cancers
CC and leukaemia, Alzheimer's disease, Parkinson's disease, multiple
CC sclerosis, epilepsy, hepatic disorders and cardiovascular disorders. This
CC sequence represents the human 46455 protein of the invention. Note: The
CC sequence data for this patent is also available in electronic format from
CC USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 528 AA;
Query Match 44.5%; Score 1298; DB 8; Length 528;
Best Local Similarity 53.4%; Pred. No. 2.4e-108;
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;
QY 53 AGDEVQTLSGSVRRAPTGP-PGTGTPGCAATAKPGGAQPKPASIGR-----GRGA-A 104
Db 2 AGSDTAPFLSQADDDPDGVPFGTGLPGSTGNPKSEEPVQDGLQRTIGLSPGRSALI 61
QY 105 AAILSLGNVNLDRYTVAGVLLDIOQHFGVDRGAGLQSVFICSFMAAPFVGLGDR 164
Db 62 VAVLCYINLLNMDRFTVAGVLPDIEQFNIGDSSGLIQTVFISYMWLAFVFGVGLDR 121
QY 165 FNRKVLSCGIPFWSAVTFSSSFIPOQYFWLLVLSRGLVIGGEASYSTIAPTIGDLFTK 224
Db 122 YNRKYLKMGCGIAFWLSVLTGSSFIPEHFWLLLTGLVGVGEASYSTIAPTIALDFVA 181
QY 225 NTRTLMVSYFAIPGLSGGLVITGSSVKQAAGDHWALRVSPVLGMITGTLILVPAT 284
Db 182 DQSRMLSFYFAIPVGSGLGYIAGSKVDMAGDHWALRVTPGLGVWAVLLFLVVRP 241
QY 285 KEGHADQLDQLKAR-TSWLRDMKALIRNRSVVFSSLATSAVSFATGALGMWIPLYLHRA 343
Db 242 PRGAVERHSDLPPLNPTSWADRLARNPFSVLSLGFATAVFTGSLALWAPAFLLRS 301
QY 344 QVVQKTAETC-NSPPCGAKDSLIFGAITCFTGFLGVVTVGAGATRCRLKQTRADPLVCAV 402
```

Db 302 RVLGETPPCLPGSDSCSSDSLIFGLITCLTGVILGVGLGVEISRRLRHSNPRADPLVCAT 361  
 Qy 403 GMLGSAIFICLI FVAAKSSIVGAYICIFVGETLLFSNWAITADILMYVVIPTTRATAVAL 462  
 Db 362 GLLGSAPFLSLACARGSI VATYIFIFIGETLLSMNWAIVADILLVVIPTTRSTAEAF 421  
 Qy 463 QSFTSHLLGDAGSPYLIGFISDLRQSTKDSPLAEFLSLGVALMCLCPVVVLGGMFFLAT 522  
 Db 422 QIVLSHLLGDAGSPYLIGLISDRURNWPPSFLSEFRALQSLMLCAFVGALGGAFLGT 481  
 Qy 523 ALFFVSDRARAEOHLGE--RRAGVVR-----VVHQRG 551  
 Db 482 AIFTEADRRRAQLHVQGLLHEAGSTDDRIVVVQRG 516

RESULT 8  
 ADL61327  
 ID ADL61327 standard; protein; 528 AA.  
 XX  
 AC ADL61327;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE Human protein tyrosine kinase biomarker spinster-like protein.  
 XX  
 KW predictor set; protein tyrosine kinase; cytostatic; antiangiogenic;  
 KW vasotrophic; vulnery; pharmacogenomic; drug sensitivity; breast cancer;  
 KW hypervascular disease; angiogenesis; wound healing scar; human;  
 KW biomarker; spinster-like.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004020583-A2.  
 XX  
 PD 11-MAR-2004.  
 XX  
 PF 26-AUG-2003; 2003WO-US026491.  
 XX  
 PR 27-AUG-2002; 2002US-0406385P.  
 XX  
 PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX  
 PI Huang F, Han X, Reeves KA, Amler L, Fairchild CR, Lee FY;  
 PI Shaw P;  
 XX  
 DR WPI; 2004-239171/22.  
 DR N-PSDB; ADL61203.  
 XX  
 PT New predictor sets with a plurality of polynucleotides and/or  
 PT polypeptides whose expression pattern predicts cell response to a  
 PT compound that modulates protein tyrosine kinase activity, useful in  
 PT treating breast cancer.  
 XX  
 PS Claim 9; SEQ ID NO 251; 649pp; English.  
 XX

CC The invention relates to a novel predictor set comprising a plurality of  
 CC polynucleotides and/or polypeptides whose expression pattern is  
 CC predictive of the response of cells to treatment with a compound that  
 CC modulates protein tyrosine kinase activity or members of the protein  
 CC tyrosine kinase pathway. The molecules of the invention demonstrate  
 CC cytostatic, antiangiogenic, vasotrophic and vulnery activities and may  
 CC be useful in the field of pharmacogenomics, in particular for determining  
 CC drug sensitivity and in treating breast cancer, hypervascular diseases,  
 CC angiogenesis and scars in wound healing. The current sequence is that of  
 CC a human protein tyrosine kinase biomarker protein of the invention.  
 XX  
 SQ Sequence 528 AA;

Query Match 44.5%; Score 1298; DB 8; Length 528;  
 Best Local Similarity 53.4%; Pred. No. 2.4e-108;  
 Matches 275; Conservative .62; Mismatches 162; Indels 16; Gaps 7;  
 Qy 53 AGDEVQTLGSGVRRAPTGP-PGTGTPGCAATAKGGAQPKPASLGR-----GRGA-A 104

Db 2 AGSDTAPFLSQADDDPGVPFGTGLPGSTGNPKSEPEVPDQEGLOKITGLSPERSALI 61  
 Qy 105 AAILSLGNLNYLDRTYVAGVLLDIOQHFGVKDRAGLQSVFICSFVAAPIFGYLGDR 164  
 Db 62 VAVLCYINLLNMYDRFTVAGVLPDIEQFNIGDSSGLIQTVFISSYVWVLPVFGYLGDR 121  
 Qy 165 FNRKVIUSCGIFPWSAVTFSSSFIPOQYFWLLVLSRGLVGIGEASYSTIAPTIGDLPFK 224  
 Db 122 YNRKYLKCGGIAFWLSVLTGLSSFPFGSHFWLLLLTRGLVGVEASYSTIAPTILADI FVA 181  
 Qy 225 NTRTLMLSVEFAIPGLSGLYITGSSVKQAAGDMHWALRVSPVLGMITGTLLIILVLPAT 284  
 Db 182 DQSRMLSIIFYAIPVSGSLGYTAGSKYKQDMAGDMHWALRVTPGLGVVAVLLFLVWREP 241  
 Qy 285 KRGHADOLGQDKAR-TSWLRDMKALIRNRSYVFFSLATSVAVSFATGALGMWIPLYLHRA 343  
 Db 242 PRGAVEHSDLPPLNPTISWADIRALARNFSVLSSLGFTAVAFVTSGLMWAFAFLRS 301  
 Qy 344 QVVKQTAETC-NSPPCGAKDSLIFGATCTFTGLGVVTVGAGATRWCLKTORADPLVCV 402  
 Db 302 RVLGETPPCLPGDSCSSDSLIFGLITCLTGVILGVGLGVEISRRLRHSNPRADPLVCAT 361  
 Qy 403 GMLGSAIFICLI FVAAKSSIVGAYICIFVGETLLFSNWAITADILMYVVIPTTRATAVAL 462  
 Db 362 GLLGSAPFLSLACARGSI VATYIFIFIGETLLSMNWAIVADILLVVIPTTRSTAEAF 421  
 Qy 463 QSFTSHLLGDAGSPYLIGFISDLRQSTKDSPLAEFLSLGVALMCLCPVVVLGGMFFLAT 522  
 Db 422 QIVLSHLLGDAGSPYLIGLISDRURNWPPSFLSEFRALQSLMLCAFVGALGGAFLGT 481  
 Qy 523 ALFFVSDRARAEOHLGE--RRAGVVR-----VVHQRG 551  
 Db 482 AIFTEADRRRAQLHVQGLLHEAGSTDDRIVVVQRG 516

RESULT 9  
 ADOS7285  
 ID ADO57285 standard; protein; 507 AA.  
 XX  
 AC ADO57285;  
 XX  
 DT 15-JUL-2004 (first entry)  
 XX  
 DE Kidney development associated protein seqid 52.  
 XX  
 KW nephrotropic; cardiant; antiinfertility; cerebroprotective;  
 KW neuroprotective; muscular; cytostatic; osteopathic; gene therapy;  
 KW kidney development; kidney disorder; developmental disorder;  
 KW circulatory disorder; hearing disorder; heart defect; infertility;  
 KW stroke; mental retardation; muscle defect; proliferative disorder;  
 KW bone defect; bone disorder; zebrafish.  
 XX  
 OS Danio rerio.  
 XX  
 PN US2004068763-A1.  
 XX  
 PD 08-APR-2004.  
 XX  
 PF 28-MAR-2003; 2003US-00403571.  
 XX  
 PR 29-MAR-2002; 2002US-0368760P.  
 XX  
 PA (HOPK/) HOPKINS N.  
 PA (GOLL/) GOLLING G.  
 PA (AMST/) AMSTERDAM A.  
 PA (SUNZ/) SUN Z.  
 XX  
 PI Hopkins N, Golling G, Amsterdam A, Sun Z;  
 XX  
 DR WPI; 2004-304692/28.  
 DR N-PSDB; ADO57284.  
 XX

PT New 459 nucleic acids and encoded polypeptides, useful for diagnosing,  
PT treating or preventing a kidney disorder in an organism, or in screening  
PT for compounds that modulate the development of an organism.  
XX  
XX Disclosure; SEQ ID NO 52; 347pp; English.  
XX  
XX The invention describes an isolated nucleic acid molecule (I) comprising  
CC a sequence having at least 75% sequence identity to the 459 nucleic acid  
CC sequence of 2808 base pairs (SEQ ID NO: 59) given in the specification,  
CC over at least 600 contiguous base pairs, where the nucleic acid functions  
CC in kidney development. (I) is useful for treating or preventing a kidney  
CC disorder in an organism, where the nucleic acid elicits an alteration in  
CC expression of a 459 nucleic acid sequence in the organism and  
CC subsequently treats or prevents a kidney disorder. The nucleic acid may  
CC also be used in diagnosing, preventing and treating a variety of  
CC mammalian diseases and developmental disorders (e.g. circulatory  
CC disorders, hearing disorders, heart defect, infertility, stroke, mental  
CC retardation, muscle defects, proliferative disorders, or bone defects or  
CC disorders) as well as in screening for compounds that modulate the  
CC development of an organism as a whole or of specific tissues or organs  
CC within that organism. This is the amino acid sequence of a kidney  
CC development associated protein.  
XX  
XX Sequence 507 AA;

Query Match 43.0%; Score 1256.5; DB 8; Length 507;  
Best Local Similarity 56.4%; Pred. No. 1.3e-104;  
Matches 251; Conservative 66; Mismatches 125; Indels 3; Gaps 2;  
QY 107 ILSGLNVLNLDYRTVAGVLLDIOQHFGVKDRGAGLLQSVFICSFMAPIFGYLGDRFN 166  
DB 54 VLYCYNLLNMDRFTVAGVLPDIEHFFGIGDGTSGLLQTVFICSYWFLAPLFGYLGDRYN 113  
QY 167 RKVILSCGIFPFSVAVTFSSFPQQYFQVLLVLSRGLVIGEASYSYTIPTIIGDLFTKNT 226  
DB 114 RKLIMCVGIFPFSVAVTFSSFPQQYFQVLLVLSRGLVIGEASYSYTIPTIIGDLFTKNT 173  
QY 227 RTMLSLVFYFAIPVSGMGYIVGSKVDTVAKDWHMALRVTPGLGLLAVFLMLVVOEPR 286  
DB 174 RTNMLSIYFAIPVSGMGYIVGSKVDTVAKDWHMALRVTPGLGLLAVFLMLVVOEPR 233  
QY 287 GHADQLGQDKARTSLWLRDMKALIRNRSVVFSSLATSVSFATGALGMWIPLYLHRAQV 346  
DB 234 GAIEAHPETHLRTSLWLRDMKALIRNRSVVFSSLATSVSFATGALGMWIPLYLHRAQV 293  
QY 347 QKTAETCNPPCGAKDLSIFGATICTPTGFLGVVGTGATWCRCLKTKQADPLVCAVGLMG 406  
DB 294 TGVKQPCFKAPCDDSDSLIFGATICTPTGFLGVVGTGATWCRCLKTKQADPLVCAVGLMG 353  
QY 407 SAIFICLI FVAAKSSIVGAYICIFVGETLLFSNMAITADILMVVITPTTRATAVALQSFT 466  
DB 354 AAPFLYLSIMFAQASTVATVYFIFLGETFLSNMAIVADILMVVITPTTRATAVALQSFT 413  
QY 467 SHLLGDAGSPYILGIFTSIDLIROSTKDSPLWEFLSLGYALMLCPFVVVLGGMFFLATLRF 526  
DB 414 SHLLGDAGSPYILGIFTSIDLIROSTKDSPLWEFLSLGYALMLCPFVVVLGGMFFLATLRF 471  
QY 527 VSDRAPEAQHGERRRAGRVVTHQRG 551  
DB 472 EKORDLAENYVPSDDAPI-VVPRSG 495

RESULT 10  
AAM40483  
ID AAM40483 standard; protein; 581 AA.  
XX  
XX AAM40483;  
XX  
XX 22-OCT-2001 (first entry)  
XX  
XX Human polypeptide SEQ ID NO 5414.  
XX  
XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
XX Homo sapiens.  
XX OS  
XX WO200153312-A1.  
XX  
XX 26-JUL-2001.  
XX  
XX 26-DEC-2000; 2000WO-US034263.  
XX  
XX 23-DEC-1999; 99US-00471275.  
XX 21-JAN-2000; 2000US-00488725.  
XX 25-APR-2000; 2000US-00552317.  
XX 20-JUN-2000; 2000US-00598042.  
XX 19-JUL-2000; 2000US-00620312.  
XX 03-AUG-2000; 2000US-00653450.  
XX 14-SEP-2000; 2000US-00662191.  
XX 19-OCT-2000; 2000US-00693036.  
XX 29-NOV-2000; 2000US-00727344.  
XX  
XX (HYSEQ-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
XX Zhou P, Goodrich R, Drmanac RT;  
XX WPI; 2001-442253/47.  
XX N-PSDB; AAI59639.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders such  
XX as central nervous system injuries.  
XX  
XX Example 2; SEQ ID NO 5414; 10078pp; English.  
XX  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
XX encoded polypeptides (AAM38642-AAM42213) with nontropic, and the  
XX immunosuppressant and cytostatic activity. The polynucleotides are useful  
XX in gene therapy. A composition containing a polypeptide or polynucleotide  
XX of the invention may be used to treat diseases of the peripheral nervous  
XX system, such as peripheral nervous injuries, peripheral neuropathy and  
XX localised neuropathies and central nervous system diseases, such as  
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
XX utilisation of the activities such as: Immune system suppression,  
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
XX assays for receptor activity, arthritis and inflammation, leukaemias and  
XX C.N.S disorders. Note: The sequence data for this patent did not form  
XX part of the printed specification

Query Match 42.8%; Score 1250; DB 4; Length 581;  
Best Local Similarity 50.9%; Pred. No. 6e-104;  
Matches 268; Conservative 63; Mismatches 152; Indels 44; Gaps 7;  
QY 8 SAAAGCAEEEADEARRRRRGAQRGAGSG-CCGARGAGGAGVSAAGDEVOTLGSVRR 66  
DB 85 TGAAGAPRDHGRVHRHRRPSSARRMTRTGTQCLAPRG- 123  
QY 67 APTGPPGT--PGTP-----GCAATAKGPAQOPKPSLGRGRGAAATLSLGNVLNLYLD 118  
DB 124 --QGPRGTRSPRSRSTRRRCGSAS-----PACLPCCSALIVAVLCINLLNYMD 171  
QY 119 RYTVAGVLLDIOQHFGVKDRGAGLLQSVFICSFMAPIFGYLGDRFNKVLSCGIFWF 178  
DB 172 RFTVAGVLPDIEHFFGIGDGTSGLLQTVFISYMWLAPVFGYLGDRYNKYLKMGCIATF 231  
QY 179 SAVTFSSSFPQQYFQVLLVLSRGLVIGEASYSYTIPTIIGDLFTKNTTLMLSVFPFAI 238

Db 232 SLVTLGSSFIPEHFWLLLTGRLVGVGEASYSYTIPTLIADLFVADQSRMLSFYFAI 291  
Qy 239 PLGSLGLYITGSSVKQAGDWHWALRSPVLGMITGTILILVPAKRGHADQLGDQLKA 298  
Db 292 PVGSLGLYIAGSKVDWAGDWHWALRVTPLGVAVVLLFLVREPRGAVRHSDDLPL 351  
Qy 299 R-TSWLRDMKALIRNRSYVSSLSATSAVSPATGALGMWIPLYLHRAQVQKTAETC-NSP 356  
Db 352 NPTSWADLRALARNPFSVLSSLGTAFAVFTGSLALWAPAFLLRSRVLGETPPCLPGD 411  
Qy 357 PCGAKDSLIFGATTCFTGFLGVVTVGAGATRCWLKTRADPLVCAVGMLSAIFICILFV 416  
Db 412 SCSSDSLIFGLITCLTGLVGLVGEISRRRLHSNPRADPLVCAVGMLSAIFLFLSLA 471  
Qy 417 AAKSSIVGAYICIFVGETLLFSNMAITADILMYVVIPTTRATVALQSFTHLILGDGSP 476  
Db 472 CARGSIATYIFIFIGETLLSMNVAIVADILLYVVIPTTRSTAEAFQIVLSHLGLDAGSP 531  
Qy 477 YLIGFISDLIRQSKDSPLWEFLSLGVALMCLCPVVLVGMFFLIATA 523  
Db 532 YLIGLISDLRLRNWPPSFLSEFRALQFSLMLCAFVGLGGAFLGTA 578

RESULT 11  
ID ABB90265  
Xx ABB90265 standard; protein; 455 AA.

AC ABB90265;  
Xx ABB90265;  
Dt 24-MAY-2002 (first entry)

Xx Human polypeptide SEQ ID NO 2641.

Xx Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
Xx anti-allergic; hepatocytotoxic; antidiabetic; antinflammatory; antitumor;  
Xx vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
Xx cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
Xx neurological disease; infection; human; secreted protein.

Xs Homo sapiens.

Xx WO200109304-A2.

Xx 29-NOV-2001.

Xx 18-MAY-2001; 2001WO-US016450.

Xx 19-MAY-2000; 2000US-020551SP.

Xx (HUMA-) HUMAN GENOME SCI INC.

Xx Birse CE, Rosen CA;

Xx WPI; 2002-122018/16.

Xx N-PSDB; ABL90674.

Xx Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
Xx prevention of neural, immune system, muscular, reproductive,  
Xx gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
Xx disorders.

Xs Claim 11; SEQ ID NO 2641; 2081pp + Sequence Listing; English.

Xx The invention relates to novel genes (ABL9449-ABL90853) and proteins  
Xx (ABB9040-ABB9044) useful for preventing, treating or ameliorating  
Xx medical conditions e.g. by protein or gene therapy. The genes are  
Xx isolated from a range of human tissues disclosed in the specification.  
Xx The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
Xx the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
Xx ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
Xx breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
Xx disorders e.g. Addison's disease, allergies, autoimmune haemolytic

CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
CC infectious diseases such as viral, bacterial, fungal and parasitic  
CC infections. Note: the sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
Xx

Seq Sequence 455 AA;

Query Match 41.8%; Score 1222; DB 5; Length 455;  
Best Local Similarity 56.7%; Pred. No. 1.5e-101;  
Matches 251; Conservative 58; Mismatches 126; Indels 8; Gaps 4;

Qy 117 LDRYTVAGVLLDQQHFGVKDRGAGLLQSVFICSEFWAAPFVGLGDRNKRKVLSCGIF 176  
Db 1 MDRFTVAGVLPDLEQEFNIGDSSGLQTVFISYMWLAPVGLGDRYNRKYLWCGGIA 60

Qy 177 FWSAVTFSSFIPOQYFVLLVLSRGLVGEASYSYTIPTLIADLFVADQSRMLSFYF 236  
Db 61 FWSLVTGLSSFIPEHFWLLLTGRLVGVGEASYSYTIPTLIADLFVADQSRMLSFYF 120

Qy 237 AIPGLSGLYITGSSVKQAGDWHWALRSPVLGMITGTILILVPAKRGHADQLGDQL 296  
Db 121 AIPVGSGLYIAGSKVDWAGDWHWALRVTPLGVAVVLLFLVREPRGAVRHSDDL 180

Qy 297 KAR-TSWLRDMKALIRNRSYVSSLSATSAVSPATGALGMWIPLYLHRAQVQKTAETC-N 354  
Db 181 PLNPTSWADLRALARNPFSVLSSLGTAFAVFTGSLALWAPAFLLRSRVLGETPPCLP 240

Qy 355 SPPCGAKDSLIFGATTCFTGFLGVVTVGAGATRCWLKTRADPLVCAVGMLSAIFICLI 414  
Db 241 GDSCSSDSLIFGLITCLTGLVGLVGEISRRRLHSNPRADPLVCAVGMLSAIFL 300

Qy 415 FVAKSSIVGAYICIFVGETLLFSNMAITADILMYVVIPTTRATVALQSFTHLILGDG 474  
Db 301 LACARGSIATYIFIFIGETLLSMNVAIVADILLYVVIPTTRSTAEAFQIVLSHLGLD 360

Qy 475 SPYLIGFISDLIRQSKDSPLWEFLSLGVALMCLCPVVLVGMFFLIATAFPVSDRAE 534  
Db 361 SPYLIGLISDLRLRNWPPSFLSEFRALQFSLMLCAFVGLGGAFLGTAIFIEADRRRA 420

Qy 535 QHLGE--RRAGVR---VVHQRG 551

Db 421 LHVQGLLHEAGSTDDRIVVVPRG 443

RESULT 12

ID AAB23610

Xx AAB23610 standard; protein; 244 AA.

AC AAB23610;

Xx 12-JAN-2001 (first entry)

Xx Human secreted protein SEQ ID NO: 20.

Xx Human; secreted protein; cytokine; cell proliferation;  
Xx nutritional supplement; immune modulation; autoimmune disorder;  
Xx haematopoiesis regulation; tissue growth; haemostasis; inflammation.

Xs Homo sapiens.

Xx Key Location/Qualifiers

FT Peptide 11..23

FT Protein 24..244

FT /label= mature\_protein

PN WO200049134-A1.

PD 24-AUG-2000.

XX 18-FEB-2000; 2000WO-US004340.  
PF 19-FEB-1999; 99US-0120680P.  
XX 23-APR-1999; 99US-00298733.  
PR 17-AUG-1999; 99US-0149639P.  
PR 23-SEP-1999; 99US-0155686P.  
PR 01-OCT-1999; 99US-0157247P.  
PR 28-NOV-1999; 99US-0167822P.  
PR 29-NOV-1999; 99US-0167823P.  
PR 15-FEB-2000; 2000US-0182711P.  
XX (ALPH-) ALPHAGENE INC.  
XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;  
XX WPI; 2000-549267/50.  
DR N-PSDB; AAA93110.  
DR New secreted proteins and polynucleotides encoding them, which are  
PT derived from Homosapiens, useful for therapy, diagnosis, and research, as  
PT well as nutritional sources or supplements.  
XX Claim 29; Page 254; 309pp; English.  
XX The present sequence is the sequence of a human secreted protein. Its  
CC cDNA was isolated from an adult pancreas cDNA library. The proteins and  
CC coding sequences of the invention can be used in the isolation of similar  
CC genes and proteins, in the elucidation of their function in vivo, and to  
CC treat a number of conditions. It is possible that they may have uses as  
CC nutritional supplements, as cytokine or cell proliferation factors, in  
CC immune modulation, where they may be used to treat immune and autoimmune  
CC diseases, as haematopoiesis regulators (treating myeloid or lymphoid cell  
CC deficiencies), in the promotion of tissue growth, they may have chemokine  
CC or chemotactic activity, haemostatic or thrombolytic activity, or anti-  
CC inflammatory activity  
XX Sequence 244 AA;  
Query Match 40.3%; Score 1176; DB 3; Length 244;  
Best Local Similarity 97.9%; Pred. No. 9.8e-98;  
Matches 229; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 306 MKALINRSVFSLSATSAVSPATGALGMWIPYLHRAQVQVKTATCNPPCGAKDSL 365  
DB 1 MKALINRSVFSLSATSAVSPATGALGMWIPYLHRAQVQVKTATCNPPCGAKDSL 60  
QY 366 FGAITCFTGLVVTGAGATRCRLKTQRADPLVCVAVGMLGSAIFCLIFVAAKSIVGA 425  
DB 61 FGAITCFTGLVVTGAGATRCRLKTQRADPLVCVAVGMLGSAIFCLIFVAAKSIVGA 120  
QY 426 YICIFVGETLLFSNWAITADILMYVVIPTERRATAVALQSFTHLLGDAGSPYLIGFISDL 485  
DB 121 YICIFVGETLLFSNWAITADILMYVVIPTERRATAVALQSFTHLLGDAGSPYLIGFISDL 180  
QY 486 IROSTKXDSLWEFLSLGYALMLCPFFVVLGSGMFFLATALFFVSDRABQHLGE 539  
DB 181 IROSTKXDSLWEFLSLGYALMLCPFFVVLGSGMFFLATALFFVSDRABQVQVQ 234  
RESULT 13  
AAE06616  
ID AAE06616 standard; protein; 476 AA.  
XX AAE06616;  
AC AAE06616;  
XX 25-SEP-2001 (first entry)  
XX Human protein having hydrophobic domain, HP10735.  
XX Human; hydrophobic domain; gene therapy; nutritional supplement;  
KW cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;  
XX multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;

KW haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;  
KW Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;  
KW haemostatic; thrombolytic; tumour growth inhibitor; anabolic;  
KW contraceptive; antiinfertility; antiinflammatory.  
OS Homo sapiens.  
XX WO200149728-A2.  
PN 12-JUL-2001.  
PD 28-DEC-2000; 2000WO-JP009359.  
XX 06-JAN-2000; 2000JP-00000585.  
PR 06-JAN-2000; 2000JP-00000588.  
PR 11-JAN-2000; 2000JP-00002299.  
PR 03-FEB-2000; 2000JP-00026862.  
PR 03-MAR-2000; 2000JP-00058367.  
XX (PROT-) PROTEGENE INC.  
PA (SAGA ) SAGAMI CHEM RES CENT.  
XX Kato S, Kimura T;  
XX WPI; 2001-418355/44.  
DR N-PSDB; AAD12611.  
XX Human proteins with hydrophobic domains and the nucleic acids encoding  
PT them, useful for preventing diagnosing and treating e.g. cancer,  
PT Alzheimer's and inflammation.  
XX Claim 1; Page 150-151; 563pp; English.  
XX The present sequence is human protein with hydrophobic domain, HP10735.  
CC The polynucleotide and polypeptide of the invention may be used in the  
CC prevention, diagnosis and treatment of diseases associated with  
CC inappropriate polypeptide expression. The polynucleotides may be used to  
CC produce the polypeptide, by inserting the nucleic acids into a host cell  
CC and culturing the cell to express the protein. The polynucleotides and  
CC its complementary sequences may also be used as DNA probes in diagnostic  
CC assays and also used in gene therapy. The polypeptides may also be used  
CC as antigens in the production of antibodies and in assays to identify  
CC modulators of polypeptide expression and activity. The polypeptides and  
CC nucleic acids may be used as nutritional supplements, to modulate  
CC cytokine and cell proliferation activity, to modulate immune stimulation  
CC or suppression (e.g. for the treatment of microbial infections and  
CC autoimmune disorders such as multiple sclerosis, rheumatoid arthritis and  
CC insulin-dependent diabetes), to modulate haematopoiesis, to modulate  
CC tissue growth activity (e.g. for the treatment of Parkinson's disease,  
CC Huntington's disease and Alzheimer's disease), to modulate activin and  
CC inhibin activity (e.g. for controlling fertility), to modulate  
CC chemotactic and chemokinetic activity, to modulate haemostatic and  
CC thrombolytic activity, to modulate receptor ligand activity, to modulate  
CC inflammation and to inhibit tumour growth  
XX Sequence 476 AA;  
Query Match 40.0%; Score 1167; DB 4; Length 476;  
Best Local Similarity 49.4%; Pred. No. 1.5e-96;  
Matches 254; Conservative 53; Mismatches 141; Indels 66; Gaps 7;  
QY 53 AGDEVQTLGSGSVRRAPTGP-PGTPTPGCAATAKGPQAQPKPASLGR-----GRGA-A 104  
DB 2 AGSDTAPFLSQADDDPDGPFVGTGTLPGSTGNPKSEEPVDPQEQLOITGLSPERSALI 61  
QY 105 AAILSLGNVLYLDRTYVAGVLLDIQQHFGVKDRAGLLQSVFICSFVWAAPIFYLGR 164  
DB 62 VAVLCYINLLNMDRFTVAGVLPDIEQPFNIGDSSGLIQTVFVFISSVNVLPAPVFGYLGR 121  
QY 165 FNRKVLSCGIFFWSAVTFSSFPQQYFWLLVLSRGLVGI GEASYSYTIPTIGDLTK 224  
DB 122 YNRKYLKMGCGIAFWSLVTLGSSFPGEHFHFWLLLRGLVGVGEASYSYTIPTIADLPFA 181





QY 524 LFFVSDRAAEQHLGE--RRAGVR-----VVHQRG 551  
Db 431 IFTEADRRRAQLHVQGLLHEAGSTDDRIIVVPQRG 464

QY 450 VVIPTRRATAVALOSFTSHLLGDAGSPYLIGFISDLIRQTKDGPLWEFLSLGYALMLCP 509  
Db 301 VVIPTRRSTAEAFQIVLSHLLGDAGSPYLIGLISDLRLRRNWPSPFLSEFRALQFSMLCA 360

QY 510 FVVVLGGMFFLTATLFFVSDRAAEQHLGE--RRAGVR-----VVHQRG 551  
Db 361 FVGALGGAAFLGTAIFTEADRRRAQLHVQGLLHEAGSTDDRIIVVPQRG 408

Search completed: October 23, 2004, 13:12:09  
Job time : 93 secs

RESULT 15

ID ADL08448 standard; protein; 420 AA.

AC ADL08448;

DT 06-MAY-2004 (first entry)

DE Human cancer suppressor protein PP2030.

KW cancer suppressor; cancer.

OS Homo sapiens.

PN CN1403480-A.

PD 19-MAR-2003.

PF 12-SEP-2001; 2001CN-00126728.

PR 12-SEP-2001; 2001CN-00126728.

PA (SHAN-) SHANGHAI XINSHIJI GENE TECHN DEV CO LTD.

PI Gu J, Yang S;

DR WPI; 2003-494229/47.

DR N-PSDB; ADL08449.

PT Human protein with function of suppressing cancer cell growth and its coding sequence.

PS Claim 1; SEQ ID NO 2; 42pp; Chinese.

CC The invention discloses one kind of human protein with cancer suppressing function, polynucleotides encoding the polypeptide and the recombinant process of producing the polypeptide. The polypeptide can be used in treating various diseases, such as cancer. The present invention also discloses the agonist resisting the polypeptide and its treatment effect. CC This sequence represents one of the cancer suppressing proteins.

XX Sequence 420 AA;

Query Match 38.2%; Score 1115; DB 7; Length 420;  
Best Local Similarity 56.6%; Pred. No. 6.4e-92;  
Matches 231; Conservative 50; Mismatches 119; Indels 8; Gaps 4;

QY 152 MVAAPIFGYLGRFRNKKVILSCGIFPWSAVTFSSSFIPOQYFWLLVLSRGLVGIGEAYS 211  
Db 1 MVLAPVFGYLDGRYRNKYLKMGCIAPWSLVTLGSSFIPEGHFHFWLLLTGLVGVGEAYS 60

QY 212 TIAPTILIGLFTKNTETMLSVFYFAIPLGSLGKGYITGSSVKQAQDWHALRVSPVLGM 271  
Db 61 TIAPTILADLFVADQKRSMLSIYFAIPVSGLVGIAGSKVDMAGDWHALRVTPGLGV 120

QY 272 ITGTILILVPATKRGHADQLGDKAR-TSWLRDMKALIRNSYVFSLSATSAVSFATG 330  
Db 121 VAVLLFLVVRPPRGAVRHSDDLPLNPTSNWADLRALARNPSFVLSLGLTAVAFVTG 180

QY 331 ALGMMIPLYLHRAQVQKTAETC-NSPPCGAKDSLIFGAITCTGTGFLGVVTCAGATRCR 389  
Db 181 SLALWAPAFLLRSRVVLGETPPCLPGDCSSDSLIFGLITCLTGLVGLGVEISRRLR 240

QY 390 LKTRADPLVCAVGLGSAIFICLIFVAAKSSIVGAYICIFVGETILFSNWAITADILMY 449  
Db 241 HSNPRADPLVCAVGLGSAIFLFLSLACARGSIATYIFIFIGETILLSNNWAIVADILLY 300

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 13:10:39 ; Search time 27 Seconds  
(without alignments)

1390.222 Million cell updates/sec

Title: US-10-085-198-48

Perfect score: 2920

Sequence: 1 MMCLCSAAGGAEEAD.....VHQRGPGTGAHRRVVGAS 566

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	779	26.7	299	4	US-09-149-476-396
2	672	23.0	424	4	US-09-149-476-555
3	396	13.6	136	4	US-09-149-476-554
4	250.5	8.6	458	4	US-09-328-352-7272
5	206.5	7.1	352	4	US-09-252-991A-30568
6	206.5	7.1	430	4	US-09-489-039A-13997
7	202.5	6.9	488	4	US-10-162-012-46
8	197.5	6.8	408	4	US-09-328-352-5768
9	189.5	6.5	443	4	US-09-489-039A-9335
10	188	6.4	506	4	US-09-252-991A-17560
11	186	6.4	463	4	US-09-489-039A-10827
12	183.5	6.3	503	4	US-09-489-039A-10722
13	182	6.2	444	4	US-09-492-709A-258
14	180	6.2	425	4	US-09-489-039A-12384
15	176.5	6.0	442	4	US-09-711-164-428
16	173.5	5.9	423	4	US-09-328-352-6097
17	173	5.9	455	4	US-09-489-039A-9942
18	173	5.9	433	4	US-09-489-039A-11835
19	168	5.8	446	4	US-09-489-039A-7920
20	168	5.8	465	4	US-09-328-352-7326
21	167	5.7	463	4	US-09-252-991A-29935
22	166.5	5.7	443	4	US-09-328-352-7567
23	166.5	5.7	460	4	US-09-252-991A-28387
24	165.5	5.7	500	4	US-09-252-991A-25053
25	165.5	5.7	502	4	US-09-489-039A-13185
26	165	5.7	415	4	US-09-489-039A-10457
27	163.5	5.6	443	4	US-09-489-039A-9381

Sequence 6461, Ap  
Sequence 9663, Ap  
Sequence 5305, Ap  
Sequence 4726, Ap  
Sequence 10203, A  
Sequence 7076, Ap  
Sequence 32216, A  
Sequence 25836, A  
Sequence 17, Appl  
Sequence 2084, Ap  
Sequence 8, Appl  
Sequence 5594, Ap  
Sequence 6006, Ap  
Sequence 11902, A  
Sequence 10630, A  
Sequence 3422, Ap  
Sequence 5706, Ap

28 163 5.6 418 4 US-09-328-352-6461  
29 162.5 5.6 460 4 US-09-489-039A-9663  
30 162 5.5 401 4 US-09-543-681A-5305  
31 162 5.5 449 4 US-09-328-352-4726  
32 161.5 5.5 476 4 US-09-489-039A-10203  
33 161 5.5 421 4 US-09-543-681A-7076  
34 160.5 5.5 537 2 US-08-647-397-2  
35 160.5 5.5 559 4 US-09-252-991A-32216  
36 160 5.5 664 4 US-09-252-991A-25836  
37 158.5 5.4 484 4 US-09-266-965-17  
38 158 5.4 379 4 US-09-710-279-2084  
39 158 5.4 848 4 US-09-575-081B-8  
40 157.5 5.4 393 3 US-09-134-001C-5594  
41 157.5 5.4 533 4 US-09-107-532A-6006  
42 157 5.4 514 4 US-09-489-039A-11902  
43 156.5 5.4 459 4 US-09-489-039A-10630  
44 156 5.3 321 4 US-09-134-000C-3422  
45 155.5 5.3 475 4 US-09-328-352-5706

## ALIGNMENTS

### RESULT 1

US-09-149-476-396  
; Sequence 396, Application US/09149476  
; Patent No. 6420526  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P1  
; CURRENT APPLICATION NUMBER: US/09/149,476  
; CURRENT FILING DATE: 1998-09-08  
; EARLIER APPLICATION NUMBER: PCT/US98/04493  
; EARLIER FILING DATE: 1998-03-06  
; EARLIER APPLICATION NUMBER: 60/040,162  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,333  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/038,621  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,626  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,334  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,336  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,163  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/047,600  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,615  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,597  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,502  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,633  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,583  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,617  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,618  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,503  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,592  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,581  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,584  
; EARLIER FILING DATE: 1997-05-23

1	EARLIER APPLICATION NUMBER: 60/047,500
2	EARLIER FILING DATE: 1997-05-23
3	EARLIER APPLICATION NUMBER: 60/047,587
4	EARLIER FILING DATE: 1997-05-23
5	EARLIER APPLICATION NUMBER: 60/047,492
6	EARLIER FILING DATE: 1997-05-23
7	EARLIER APPLICATION NUMBER: 60/047,598
8	EARLIER FILING DATE: 1997-05-23
9	EARLIER APPLICATION NUMBER: 60/047,613
10	EARLIER FILING DATE: 1997-05-23
11	EARLIER APPLICATION NUMBER: 60/047,582
12	EARLIER FILING DATE: 1997-05-23
13	EARLIER APPLICATION NUMBER: 60/047,596
14	EARLIER FILING DATE: 1997-05-23
15	EARLIER APPLICATION NUMBER: 60/047,612
16	EARLIER FILING DATE: 1997-05-23
17	EARLIER APPLICATION NUMBER: 60/047,632
18	EARLIER FILING DATE: 1997-05-23
19	EARLIER APPLICATION NUMBER: 60/047,601
20	EARLIER FILING DATE: 1997-05-23
21	EARLIER APPLICATION NUMBER: 60/043,580
22	EARLIER FILING DATE: 1997-04-11
23	EARLIER APPLICATION NUMBER: 60/043,568
24	EARLIER FILING DATE: 1997-04-11
25	EARLIER APPLICATION NUMBER: 60/043,314
26	EARLIER FILING DATE: 1997-04-11
27	EARLIER APPLICATION NUMBER: 60/043,569
28	EARLIER FILING DATE: 1997-04-11
29	EARLIER APPLICATION NUMBER: 60/043,311
30	EARLIER FILING DATE: 1997-04-11
31	EARLIER APPLICATION NUMBER: 60/043,671
32	EARLIER FILING DATE: 1997-04-11
33	EARLIER APPLICATION NUMBER: 60/043,674
34	EARLIER FILING DATE: 1997-04-11
35	EARLIER APPLICATION NUMBER: 60/043,669
36	EARLIER FILING DATE: 1997-04-11
37	EARLIER APPLICATION NUMBER: 60/043,312
38	EARLIER FILING DATE: 1997-04-11
39	EARLIER APPLICATION NUMBER: 60/043,313
40	EARLIER FILING DATE: 1997-04-11
41	EARLIER APPLICATION NUMBER: 60/043,672
42	EARLIER FILING DATE: 1997-04-11
43	EARLIER APPLICATION NUMBER: 60/043,315
44	EARLIER FILING DATE: 1997-04-11
45	EARLIER APPLICATION NUMBER: 60/048,974
46	EARLIER FILING DATE: 1997-06-06
47	EARLIER APPLICATION NUMBER: 60/056,886
48	EARLIER FILING DATE: 1997-08-22
49	EARLIER APPLICATION NUMBER: 60/056,877
50	EARLIER FILING DATE: 1997-08-22
51	EARLIER APPLICATION NUMBER: 60/056,889
52	EARLIER FILING DATE: 1997-08-22
53	EARLIER APPLICATION NUMBER: 60/056,893
54	EARLIER FILING DATE: 1997-08-22
55	EARLIER APPLICATION NUMBER: 60/056,630
56	EARLIER FILING DATE: 1997-08-22
57	EARLIER APPLICATION NUMBER: 60/056,878
58	EARLIER FILING DATE: 1997-08-22
59	EARLIER APPLICATION NUMBER: 60/056,662
60	EARLIER FILING DATE: 1997-08-22
61	EARLIER APPLICATION NUMBER: 60/056,872
62	EARLIER FILING DATE: 1997-08-22
63	EARLIER APPLICATION NUMBER: 60/056,882
64	EARLIER FILING DATE: 1997-08-22
65	EARLIER APPLICATION NUMBER: 60/056,637
66	EARLIER FILING DATE: 1997-08-22
67	EARLIER APPLICATION NUMBER: 60/056,903
68	EARLIER FILING DATE: 1997-08-22
69	EARLIER APPLICATION NUMBER: 60/056,888
70	EARLIER FILING DATE: 1997-08-22
71	EARLIER APPLICATION NUMBER: 60/056,879
72	EARLIER FILING DATE: 1997-08-22
73	EARLIER APPLICATION NUMBER: 60/056,880

1	EARLIER FILING DATE: 1997-08-22	60/056,894
2	EARLIER APPLICATION NUMBER: 60/056,894	
3	EARLIER FILING DATE: 1997-08-22	60/056,911
4	EARLIER APPLICATION NUMBER: 60/056,911	
5	EARLIER FILING DATE: 1997-08-22	60/056,636
6	EARLIER APPLICATION NUMBER: 60/056,636	
7	EARLIER FILING DATE: 1997-08-22	60/056,874
8	EARLIER APPLICATION NUMBER: 60/056,874	
9	EARLIER FILING DATE: 1997-08-22	60/056,910
10	EARLIER APPLICATION NUMBER: 60/056,910	
11	EARLIER FILING DATE: 1997-08-22	60/056,864
12	EARLIER APPLICATION NUMBER: 60/056,864	
13	EARLIER FILING DATE: 1997-08-22	60/056,631
14	EARLIER APPLICATION NUMBER: 60/056,631	
15	EARLIER FILING DATE: 1997-08-22	60/056,845
16	EARLIER APPLICATION NUMBER: 60/056,845	
17	EARLIER FILING DATE: 1997-08-22	60/056,892
18	EARLIER APPLICATION NUMBER: 60/056,892	
19	EARLIER FILING DATE: 1997-08-22	60/057,761
20	EARLIER APPLICATION NUMBER: 60/057,761	
21	EARLIER FILING DATE: 1997-08-22	60/047,595
22	EARLIER APPLICATION NUMBER: 60/047,595	
23	EARLIER FILING DATE: 1997-05-23	60/047,599
24	EARLIER APPLICATION NUMBER: 60/047,599	
25	EARLIER FILING DATE: 1997-05-23	60/047,588
26	EARLIER APPLICATION NUMBER: 60/047,588	
27	EARLIER FILING DATE: 1997-05-23	60/047,585
28	EARLIER APPLICATION NUMBER: 60/047,585	
29	EARLIER FILING DATE: 1997-05-23	60/047,586
30	EARLIER APPLICATION NUMBER: 60/047,586	
31	EARLIER FILING DATE: 1997-05-23	60/047,590
32	EARLIER APPLICATION NUMBER: 60/047,590	
33	EARLIER FILING DATE: 1997-05-23	60/047,594
34	EARLIER APPLICATION NUMBER: 60/047,594	
35	EARLIER FILING DATE: 1997-05-23	60/047,589
36	EARLIER APPLICATION NUMBER: 60/047,589	
37	EARLIER FILING DATE: 1997-05-23	60/047,593
38	EARLIER APPLICATION NUMBER: 60/047,593	
39	EARLIER FILING DATE: 1997-05-23	60/047,614
40	EARLIER APPLICATION NUMBER: 60/047,614	
41	EARLIER FILING DATE: 1997-05-23	60/043,578
42	EARLIER APPLICATION NUMBER: 60/043,578	
43	EARLIER FILING DATE: 1997-04-11	60/043,576
44	EARLIER APPLICATION NUMBER: 60/043,576	
45	EARLIER FILING DATE: 1997-04-11	60/047,501
46	EARLIER APPLICATION NUMBER: 60/047,501	
47	EARLIER FILING DATE: 1997-05-23	60/043,670
48	EARLIER APPLICATION NUMBER: 60/043,670	
49	EARLIER FILING DATE: 1997-04-11	60/056,632
50	EARLIER APPLICATION NUMBER: 60/056,632	
51	EARLIER FILING DATE: 1997-08-22	60/056,664
52	EARLIER APPLICATION NUMBER: 60/056,664	
53	EARLIER FILING DATE: 1997-08-22	60/056,876
54	EARLIER APPLICATION NUMBER: 60/056,876	
55	EARLIER FILING DATE: 1997-08-22	60/056,881
56	EARLIER APPLICATION NUMBER: 60/056,881	
57	EARLIER FILING DATE: 1997-08-22	60/056,909
58	EARLIER APPLICATION NUMBER: 60/056,909	
59	EARLIER FILING DATE: 1997-08-22	60/056,908
60	EARLIER APPLICATION NUMBER: 60/056,908	
61	EARLIER FILING DATE: 1997-08-22	60/048,964
62	EARLIER APPLICATION NUMBER: 60/048,964	
63	EARLIER FILING DATE: 1997-06-06	60/057,650
64	EARLIER APPLICATION NUMBER: 60/057,650	
65	EARLIER FILING DATE: 1997-09-05	60/056,894
66	EARLIER APPLICATION NUMBER: 60/056,894	
67	EARLIER FILING DATE: 1997-08-22	

```
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match      26.7%; Score 779; DB 4; Length 299;
Best Local Similarity 56.1%; Pred. No. 1.5e-61;
Matches 162; Conservative 34; Mismatches 91; Indels 2; Gaps 2;

QY 230 MLSVFYFAIPLGSLGYITSSVKQAGDWMHWRVSPVGMITGTILILVPAIKRGHA 289
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MLSIFFAIPVSGSLGYIAGSKYKQAGDWMHWRVTPGIGVAVLLFLVPRPGAV 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 290 DQGGDQLKAR-TSWLRDMKALINRSYVFSLSATSAVSFATGALGMWIPLYLHRAOVQK 348
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 ERHSDLPPLNPTSWADRLARALNPSPVLSLGTAVAFVTSGLAWAPAFILRSRVLG 120
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 349 TATC-NSPPCGAKDSLIFGAITCTGFLGVTVTGAGATRCWLKTRADPLVCAVGMG 407
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 ETPCLPGDSCSSDSLIFGLITCLTGLVGLGVEISRRLHNSPRADPLVCATGLGS 180
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 408 AIFCLIFVAAKSSIVGAYICIFVGTLLFSNWAITADILMYVITPRATVALQSFTS 467
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 APFLSLACARSGIVATYIFIGETLLSMNWAIVADILLYVITPRSTAEAFQIVLS 240
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 468 HLLGDAGSPYLIGFISDLIRQSKDQPLWFLSGLVAMLCPPFVVLGG 516
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 HLLGDAGSPYLIGLISDLRLRNWPPFLSEFRALQFSLMLCAFVGLGG 289
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 2
US-09-149-476-555
; Sequence 555, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
```

RES001 3  
US-09-149-476-554  
Sequence 554, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
CURRENT FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23



EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 13.6%; Score 396; DB 4; Length 136;  
Best Local Similarity 65.2%; Pred. No. 9.5e-28;  
Matches 73; Conservative 19; Mismatches 20; Indels 0; Gaps 0;

QY 117 LDRYTVAGVLLDIOQHFQKVRGAGLQSVFVCSFMAPIFGYGLDRNRKVLSCGIF 176  
DB 1 MDRFTVAGVLPDIOEFNGDSSGLIQTVFSSFWLAPVFGYGLDRNRKVLSCGIIA 60  
QY 177 FMSVTFSSFPQQYFWLLVLSRGLVGIGEASYSYTIPTIIGDLFTKNT 228  
DB 61 FWSLVTLGSSFPGEHFHLLLRGLVGVGEASYSYTIPTIADLFLVADORT 112

## RESULT 4

US-09-328-352-7272  
Sequence 7272, Application US/09328352  
Patent No. 6562958

## GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 7272  
LENGTH: 458  
TYPE: PRT

ORGANISM: Acinetobacter baumannii

US-09-328-352-7272

Query Match 8.6%; Score 250.5; DB 4; Length 458;  
Best Local Similarity 24.2%; Pred. No. 5.3e-14;  
Matches 111; Conservative 76; Mismatches 194; Indels 77; Gaps 17;

QY 110 LGNVLVNDRYTVAGVLLDIOQHFQKVRGAGLQSV-FTCSFMAAPIFGYGLDRNRK 168  
DB 37 LAVIFSDIRQIUALMIEPIKADLQSDQFSLHGLASLFAVVMGLPLAYIADFSRP 96  
QY 169 VILSCGIFFFWSAVT----FSSSFIPOQYFWLLVLSRGLVGIGEASYSYTIPTIIGDLFTK 224  
DB 97 KLISIGIIVWSLATATCGLSKFIQ-----LFLSRMVGVGSAALSPAAYSFMFSK 150  
QY 225 NRTMLSVFYPFAIPGLSGLYTGSVKQ-----AAGDHWALRVSPVLG 270  
DB 151 DKLGRAVGIYSIGAFLLGGGAFIPLVGGYVNNLLKGVTLIEVPLGALKAKQIAFLVGLPG 210  
QY 271 MITGTILILV--PATKRGHADQLG--DQLKARTSMLRDMKALIRNRSYVFSLSATSAVS 326

DB 211 IIGLLFILVTKOPARKGQQLNQSDQVDF--TQCLQFIKK--HAKTFACHYGLGFTFYA 267  
QY 327 FATGALGMWIP-LYLHRAQVVQKTAETCNPPCAKDSLFGAICTGTGFLGVVVTGAGAT 385  
DB 268 MALYSLSWTPAFYIRKFLAP-----TETGYMLGTILLVANTLGVFCAGWLN 315  
QY 386 RWCLKTKQRADPLVCV-GMLGSAIFICLIFVAAKSSIVGAYICIFVGETLL-----PSN 439  
DB 316 DMFIKGRQDAPMFTGVIGV--LIPIAFTTQDQ-----LWLSVILLIIPAMFPAS 366  
QY 440 W--AITADILMYVVIPTERRATAVALQSFTHLLGDAGSPYILGIFISDLIROSTKDSPLWE 497  
DB 367 FPLVISATALQMLAPNQFRARLSALFLVSNLIGLVGTTLVAITDKV-----415  
QY 498 FLSLGYALMLCPFVVVLGGM--FFLATALFFVSDRARAE 534  
DB 416 ---FGNPLMVGSSLSIVGGLSCVLALALLFKGCKSFSE 450

## RESULT 5

US-09-252-991A-30568  
Sequence 30568, Application US/09252991A  
Patent No. 6551795

## GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252.991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 30568  
LENGTH: 352  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30568

Query Match 7.1%; Score 206.5; DB 4; Length 352;  
Best Local Similarity 23.7%; Pred. No. 3.2e-10;  
Matches 87; Conservative 56; Mismatches 141; Indels 83; Gaps 14;

QY 194 MLLVLSRGLVGIGEASYSYTIPTIIGDLFTKNTRLMLSVFYFAIPLGSLGYITGSSV- 252  
DB 1 WQFLTRFVGVGVGEAALSPAAYSLSIADSPFRERRATAISVYSMGIYLSGLAFLGLGLVI 60  
QY 253 --KQAGDWH-----WAL--RVSPVLGMITGTILILVPAKRGHADQLGDLQKAR 299  
DB 61 KFAAQGDVHLPLFGEVTPWQLIFILGNAAGVLFCLLLLAIREPARRGVGAGVAVPLGEV 120  
QY 300 TSWLRDMKALIRNRSYVFSLSATSAVSFATGALGMWIPLYLHRAQVVQKTAETCNSPPCG 359  
DB 121 GAYLRANKTVLCHNFGFACL-----SPAGYGSAGAVPTFFVTRTHGWD-----A 164  
QY 360 AKDSLIFGAICTCTGFLGVVTCAG--ATRCRLKTKTORADPLVCVAGMLGSAIFICLIFVAA 418  
DB 165 GHGVVYGSIVAVFGCLGIVFGRLADYWA--KRGSDANM--RVGLL--AAWAVIFP--- 216  
QY 419 KSSIVGAYICIFVGTETLLF-----SNW--AITADILMYVVIP-----TRRA 457  
DB 217 -----TLVYPLDNNANWAAALMAPTVFFLSMPFGVAPAAIQEIMPNSMRG 261  
QY 458 TAVALQSFTHLLGDAGSPYILGIFISDLIROSTKDSPLWEFLSLGLVALMLCPFVVVLGGM 517  
DB 262 QASATYLFVVTILFGLGLGPTAVALVTFV-----FADDMALRSLLLVTLAAVLGAV 313  
QY 518 FLATLAL 524  
DB 314 VLLIGL 320



## RESULT 6

US-09-489-039A-13997  
; Sequence 13997, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489.039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 13997  
; LENGTH: 430  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-13997

Query Match 7.18; Score 206.5; DB 4; Length 430;

Best Local Similarity 22.98; Pred. No. 4.2e-10;  
Matches 102; Conservative 82; Mismatches 200; Indels 61; Gaps 17;

QY 94 PASLGRGGRGAAATLSLGNVLYLDRTVAGVLLDIOQHFGVKGDRGAGLLQSFVICSFMV 153  
DB 17 PLRIRRVQKITLMTLFTAGVFNLDSSLSVAGEAIRADIGLSNTEFGVLLSAFSLSYGF 76  
QY 154 AAPFGYLGDRFNKRLGSCIGFWSAVTFSSSFIPOQYFLLVLSGLVIGGEASVSTI 213  
DB 77 AQLPSGILLDLRLGPRIVLGLAGLIFWSAQALTCMV-NSFSHFILLRIGL-GIGEAAPMPA 134  
QY 214 APTIIGDLFTYNTLMLSVFYRAIPGSLGLYTGSSVQKQAGDWALRVSPVLGMIT 273  
DB 135 GVKSSINDWYQQRGTGAVGFNSSTVLGQAI-PPALVIMQLAWGRTFMFVIGLAGIV 193  
QY 274 GTLILILVPATKGHADQ-----LGDLQKAR-----TSWLRDMKALIRNSYV 316  
DB 194 GLCWYV-----GYRNRQFTQEBEQVLAEEAARPAKFSWL-----ALFKRTTW 242  
QY 317 FSSLATSAVSFATGALGMWIPLYLHRAQVQVQKTAETCNPPCGAKDSLIFGALITCTGFL 376  
DB 243 GMILGFGVNYTWGLYIAWLPGLV-LQAQGLSLART-----GWVAIPFLAAAGWVN 295  
QY 377 GVTGAGATR-WCKLTKQRADPLVCVAGMLGSAIFICLIFVAAKSSIVGAYI-----CIF 430  
DB 296 GLVVDALARRGYDQAKTRKT---AIVIGLVLSALGTLVVQSSSTPAQAVAFISNALFCVH 352  
QY 431 VGETLLFSNWAITADILMYVVIPTRRATAVALQSFTSHLAGDAGSPVLIGFISDLIRQST 490  
DB 353 FAGT---SAGW-----LVQVWAEHKVASVAIQNFGSFVFA-SFAPIVTGWVVD-----TT 400  
QY 491 KDSPLMEFLSLG--YALMLCPFFVW 513  
DB 401 HSFNLALVIRAGTVFAGALCYFFIV 425

## RESULT 7

US-10-162-012-46  
; Sequence 46, Application US/10162012  
; Patent No. 6682597  
; GENERAL INFORMATION:  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Silos-Santiago, Immaculada  
; APPLICANT: Gu, Wei  
; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS  
; FILE REFERENCE: 10448-130001  
; CURRENT APPLICATION NUMBER: US/10/162.012  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: US 60/209,845  
; PRIOR FILING DATE: 2000-06-06

; PRIOR APPLICATION NUMBER: US 09/875,321  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: PCT/US01/18340  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/209,257  
; PRIOR FILING DATE: 2000-06-05  
; PRIOR APPLICATION NUMBER: US 09/875,423  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: PCT/US01/18398  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/209,238  
; PRIOR FILING DATE: 2000-06-05  
; PRIOR APPLICATION NUMBER: US 09/875,363  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: PCT/US01/18247  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/227,068  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: US 09/928,530  
; PRIOR FILING DATE: 2001-08-13  
; PRIOR APPLICATION NUMBER: PCT/US01/25475  
; PRIOR FILING DATE: 2001-08-15  
; PRIOR APPLICATION NUMBER: US 60/226,770  
; PRIOR FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: US 09/934,421  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/US01/26096  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/279,281  
; PRIOR FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: US 10/109,029  
; PRIOR FILING DATE: 2002-03-28  
; PRIOR APPLICATION NUMBER: PCT/US02/09728  
; PRIOR FILING DATE: 2002-03-28  
; PRIOR APPLICATION NUMBER: US 60/290,288  
; PRIOR FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: US (not assigned)  
; PRIOR FILING DATE: 2002-05-13  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 46  
; LENGTH: 488  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: consensus sequence  
US-10-162-012-46

Query Match 6.9%; Score 202.5; DB 4; Length 488;  
Best Local Similarity 21.5%; Pred. No. 1.2e-09;  
Matches 114; Conservative 88; Mismatches 195; Indels 133; Gaps 22;  
QY 95 ASLGRC-----RGAATAILSLGNVLYLDRTVAGVLLDIOQHFGVKGDRGAGLLQS 145  
DB 5 AALGGGFLFGYDGTGVIGGFLAL---IDFLPRFGLLTSSGALAE---VGYSTVLTLGLWS 58  
QY 146 VFICSFVWAAPIFGYLGDRENK---VILSCGIFWMSAVTFSS-----SFIPQOYFWLLVLS 199  
DB 59 IFPLGLGLSLFAGLGRDRFRKKSLIALVLVEIGALLSGAAPGYTTIGLWAFYLLIV 118  
QY 200 RGLVGIGEASYSYTIATIGDLTKTRTLMLSVFYFAIPGLSGGLGYITGSSVKQAAGD- 258  
DB 119 RVLVGLGVGASVJVPWYISETAPKALRGALGSLYQLAITIGILVAAIIGLKNKTNDS 178  
QY 259 ----WHW-----ALRVSPVLGMITGTILILVP-----ATKRG--HADQL 292  
DB 179 ALNSWGWRIPLGLQLVPALLLGLLFLPESPRWLVKLEBEAREVLAKLRGVEDVDQE 238  
QY 293 GDOLKA-----RTSWLRDMKALIRN---RSYVSSLSATSATGALGMWIP 338  
DB 239 IOEIKAELEATVSEKAGKAGMGELEFRGRTPKVRQRLMLGMVLMQAFQQLTG-----INA 293  
QY 339 YLHRAQVQVQKTAETCNPPCGAKDSLIFGALITCTGFLGWV----- 379

Db 294 IYYSPTIKSV-----GVSDSVASLLVTLIIVGVNVFTFVALIFLVDRFGRPL 345  
Qy 380 -TGAGATRCRLKTORADPLVCAVGMGSAIFICLIFV-----AKSIVGAYICIFVG 432  
Db 346 LLGAGMAICFL-----ILGASIGVALLLNKPKDPSSKAAGIVAIVFILLF 392  
Qy 433 ETLTFSNWAITADILMYVVIPT-RATAVALOSFTSHLLGDAGSPVLGFIQSLIRQSTK 491  
Db 393 IAFPALGWGPIPWILSELFFPKVRSKALATAANWL-----ANFIIGFLFPYITG--- 444  
Qy 492 DSQLBEFLSLGVALMCPFWVLGGM-FFLATALFVSD-RARAEOHLE 539  
Db 445 -----AIGLALGGVFLVAGLLVLFLFVFFVFPETKRTLEIEE 486

RESULT 8  
US-09-328-352-5768  
; Sequence 5768, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 5768  
; LENGTH: 408  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-5768

Query Match 6.8%; Score 197.5; DB 4; Length 408;  
Best Local Similarity 21.6%; Pred. No. 2.5e-09;  
Matches 99; Conservative 71; Mismatches 174; Indels 115; Gaps 17;  
Qy 105 AAILSGNLVLDRTYVAGVLD--IQHFGVKDGA--GLQSVFICSPFWAIPFY 160  
Db 16 AFVLMVGVVLAFFDKISIAVLFSDFPHQSSMGIAEDKAKGLWMTSFLLAYGFSVFLSF 75  
Qy 161 LGDRFNKRVILSCGIFFWSAVTFSSSFIPOQYFWLLVLSRGLVGIGEASYSYTIPTIGD 220  
Db 76 LGDIFNPKMFLWSVTSWGLMLCMCF-TTSYSGMLIL-RVLGLAEGPLFALATYIVKQ 133  
Qy 221 LFTKTRTLMLSVYFPAIFLGSGLGYITGSSVKQAAGDWHWALRVSPLVGMITGTLLIL 280  
Db 134 TVTDROQARASTWFLIGTIGAGFLGPPIAAV-LAHHDWHTTFFVMAALTIALILSIVFG 192  
Qy 281 VPATKRGHADQLGDQKARTSW---LRDMKALIRNSYVFSSLATSAVSFATGALGMWIP 337  
Db 193 LRNLQKTKTVELEGESK-RTNFKGHANTKVLVNSAFWLVCLEFNIALTYLWGLNSWVP 251  
Qy 338 LYLHRAQVQKTAETCNCSPPCGAKOSLIFFGATCTFTGLGVVTVGAGATRCRLKTORADP 397  
Db 252 SYLMQDK-----GF-----NLK----- 263  
Qy 398 LVCVAGMLGSAIFICLIFVAAKSSIVGAYICIFVGETLL--FSNWAITADILMYV-VIPT 454  
Db 264 ---EFGVYSFFPIAMLI---GEVVGAFLSKLGRRAIQVFSG-LLLAGIFMYVMVIMT 315  
Qy 455 RRATAVALOSF-----TSHLLGDAG-----SPVLICF 481  
Db 316 EPLLIIAAMSLSAMWFGVAVFALLARVTSNMGATAGGIFNGLGNFASAIAPVLIGY 375  
Qy 482 ISDLIRQSTKDSPLMBFLSLGVALMCPFWVLGGMFFL 520  
Db 376 I-----VMQTHSFNLGITLAAVAVIGSLFLV 402

RESULT 9  
US-09-489-039A-9335

; Sequence 9335, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 9335  
; LENGTH: 443  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-9335

Query Match 6.5%; Score 189.5; DB 4; Length 443;  
Best Local Similarity 22.0%; Pred. No. 1.5e-08;  
Matches 101; Conservative 72; Mismatches 200; Indels 87; Gaps 19;  
Qy 114 LNYLDRTYVAGVLDIQHFGVKORAGAGLQSVFICSPFWAIPFYGLDRFNKRVILSC 173  
Db 29 VNYGDRATLSIAGTEVAKELGLSAVSMGYIFSAGWAYLLMQIPGGWLLDKFGSKKVYSY 88  
Qy 174 GIFPWSAVTFSSSFI---PQYFWL-LVLSRGLVGIGEASYSYTIPTIGDILFTKNTRTL 229  
Db 89 SLFPWSLFTFLOGFDVFPFLAWAGVSMFFMFLGFSFSEAPFANARIVAAMPFAKEROT 148  
Qy 230 MLSVF---YFAIPLGSG-LGYITGSSVKQAAGDWHWALRVSPLVGMITGTLLILVLPAT 284  
Db 149 ASAIFNAQAQYFSLALFSPLLGLMT-----FALGWEHVFTVWGIIGFVL-TIIVKFEVN 201  
Qy 285 KRGHADQLGDQK-----ARTSWLRDMKALIRNSYVFSSLATSAV 325  
Db 202 PTDPRMSAAELKVISEGGAVVDMHKKEATPAAGPKMDYIRQLLTNRMLGCVFVGQYPL 261  
Qy 326 SFATGALGMWIPVYL--HRAQVQKTAETCNCSPPCGAKOSLIFFGATCTFTGLGVVTVGAG 383  
Db 262 NTITWFFLTWTFPIVLQDKGWSILKVGVSAP-----ALFG---FAG---GVLGGLF 308  
Qy 384 AT---RWCRLKTORADPLVCAVGMGSAIFICLIFVAAKSSIVGAYICIFVGETLLFSN 439  
Db 309 SDYLIGRGCTLTFAKPLIVLGM-LLASSIILC-NYTASTPIVITLMALAPFGKGFALG 366  
Qy 440 WAITADILMYVVIPTRRATAVALOSFTSHLLGDAGS---PYLIGFISDLIROSTKDSPLW 496  
Db 367 WPVISDVAPKEI-----VGLCGGVNFVFNVASIATPLVIGYI-----VS 406  
Qy 497 EFLSLGVALMCPFWVLGSG--MPFLATALFVSDRAAE 534  
Db 407 ELHSPNGAL-----IFVGGSAALMMWCVLFVVGDIKWE 440

RESULT 10  
US-09-252-991A-17560  
; Sequence 17560, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 17560  
; LENGTH: 506

; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17560

Query Match 6.4%; Score 188; DB 4; Length 506;  
Best Local Similarity 20.1%; Pred. No. 2.4e-08;  
Matches 121; Conservative 72; Mismatches 218; Indels 192; Gaps 23;

QY 25 RRR----RGAQRGAGGCCGARGAGGAGVSAAGDEVQLTSGSVRRAPTGPPTGPGFC 80  
DB 4 RRRAGFIRARRPAPASQ--SARRRVAARQARRRRSOTRSTPAARQNWMPDAPAS-- 58  
QY 81 AATKAGCAQPKPASLGRGRGAAAAIISLGNVLNLYDRYTVAGVLDLIDQHFQVKDORGA 140  
DB 59 -----AABRLRLPLSPYHRLVFVIAIAFFPDSMDLAMTFLIGSKIAEFLGDSAQA 111  
QY 141 GLLOSVPICFPMVAAPFGVLDGRFNKVLSCGIFFWSAVTSSSIPQOYFWLLVLSR 200  
DB 112 GLIASSFFGKVIAGALUSGMLADRFGRKPVQASIVLWGLASVLCSTAGD--LDSLTFYR 169  
QY 201 GLVGIGEASYSTIAPTIIIGDLFTKNTRTLMLSVF-----YFAIPLGSLG 245  
DB 170 VLLGIGMGEPPIAQSLLSEMPASRRGKYIALMDGFPLGFVAAGCLSYFLPL----- 224  
QY 246 YITGSSVKQAQDMHWALRVSPVLGMITGTLLILVLP-----ATKRGHADQLGDLK 297  
DB 225 --TG-----WR-SIFLVLPALPAVFVLAIRFLIPESPRWLEQAGRRREQADR----- 267  
QY 298 ARTSWLRDMKALI-----RNRSVFSSSLATSATSPATGALGMWPL 338  
DB 268 -----LNDIEARVNRSLGTELPLPLQOPRERSRPGFFSAFAB-----LWSPA 311  
QY 339 YLHRAQVQKTAETCNSPPCGAKDSLIFGAICTGFLGVVVTGAGATRCWLKTRADPL 398  
DB 312 YRRTQTW-----GLWFFAL---LGFYGL-----TSWLSALLQSGGFA 347  
QY 399 V-----CAVGMGLGSAIFC-----LIFVAAKSSIVGAYIC 428  
DB 348 VTQSVYTVLISLAGIPGLCAAWLVBESWGKPSCVLMLLGGGAMAYAGTAVFGGSLA 407  
QY 429 IFVG-----ETLFSNWAITADILMYVIP-----TRATAVALQSTSHLLGDAGSPYL 478  
DB 408 LLIGFGLAMOFFLFGMWAV-----LYTTPELYTSARATSGSFASAVGR-IGSLGLPLV 461  
QY 479 IGFISDLIRQSTKDSPLWEFLSLGALMCPFFVVVLGGMFFLATAL-----FFVSDRARAE 534  
DB 462 TGLVLPITGCGG-----VFTLIGALCFGVALVWVAFGIETRGRTL 501  
QY 535 QHL 537  
DB 502 EEL 504

RESULT 11  
US-09-489-039A-10827  
; Sequence 10827, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCES: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 10827  
; LENGTH: 463  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-10827

Query Match 6.4%; Score 186; DB 4; Length 463;  
Best Local Similarity 21.9%; Pred. No. 3.2e-08;  
Matches 82; Conservative 68; Mismatches 165; Indels 60; Gaps 13;

QY 118 DRYTVAGVLDLIDQHFQVKDORGAGLLOSVPICFPMVAAPFGVLDGRFNKVLSCGIFP 177  
DB 50 DTAIGFIASDLVQEWGVEKSAALGPVNSAALVGLAVGALTAGPLADRIGRKKVLMISIV 109  
QY 178 WSAVTSSSIPQOYFWLLVLSRGLVGIGEASYSTIAPTIIIGDLFTKNTRTLMLSVFYFA 237  
DB 110 PGFSLTATATS--LNQLTLLFLFTGLGGAAMPNATLMSYAPERRALLVNLMPVG 167  
QY 238 IPIGSGIGYITGSSVKQAQDMHWALRVSPVLGMITGTLLILVLP-----ATKRGHADQ 291  
DB 168 FPMGSSLGGLSAAWIPHYG-WQSVLVLGVMPLLLAVALIFLLPESVRYLVVWQHPAQ 226  
QY 292 LGDQLKARTSWLRDMKALIRNSRYVSSSLATSAV---SPATGA-----L 332  
DB 227 IAAILRRIAPLPAVEFVLRAGQVKEKSAIGVIFSPRYAVGTVMCLTYFMGLLIIFYLL 286  
QY 333 GMWIPLYLHR--AOWVQKTAETCNSPPCGAKDSLIFGAICTGFLGVVVTGAGATRCWL 390  
DB 287 TSWLPLIRRETGASMSQASIIITALFPLGGGIGVLLGL----- 325  
QY 391 KTORADP-LYCAVGMGLGSAIFCII-FVAAKSSIVGAYICIFVGETLLFSNWAITA-DIL 447  
DB 326 -MDKINPNKVVAVGWLITGVFVCLVGFSTSSSLALMG--VWVFIAGSIM--NGAQSSMPAL 380  
QY 448 MYVVIPTR-RATAVA 461  
DB 381 AAGFYPTQGRATGVA 395

RESULT 12  
US-09-489-039A-10722  
; Sequence 10722, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCES: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 10722  
; LENGTH: 503  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-10722

Query Match 6.3%; Score 183.5; DB 4; Length 503;  
Best Local Similarity 22.1%; Pred. No. 6e-08;  
Matches 105; Conservative 62; Mismatches 204; Indels 105; Gaps 17;

QY 108 LSLGNVLNLYDRYTVAGVLDLIDQHFQVKDORGAGLLOSVPICFPMVAAPFGVLDGRPNR 167  
DB 26 LLLWMLLSALDQDTIVALTPTIVGELGDLKLSWV-TAYILSSTIAVPLYGRKGDIFGR 84  
QY 168 KVL--SCGIFFWSAVTSSSIPQOYFWLLVLSRGLVGIGEASYSTIAPTIIIGDLFTKN 225  
DB 85 KVLQVAIGLFLVGSALCGLA-----QNWTLVLMRGLGGLGGGLVMSAAVADVIPPA 140  
QY 226 TRTMLSVFYFAIPGLSGIGYITGSSVKQAQDMHWALRVSPVLGMITGTLLILVPAWK 285  
DB 141 NRGRYOGLFGGVFLGATVIGPLIGGLVQHA-SRWIFVYNLPLGLFALLVIGAVFHSN 199  
QY 286 RGHADQLGDLKARTSWLRDMKALIRNSRYVSSSLATSATSPAT---GALGMWIPLYLHR 342  
DB 200 KRSQHQI-----DWL-----GAIVLSMALLCILFTSREGSVHAW----- 234



```
; FILE REFERENCE: ELITRA.008A
; CURRENT APPLICATION NUMBER: US/09/711.164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-711-164-428

Query Match      6.0%; Score 176.5; DB 4; Length 425;
Best Local Similarity 22.7%; Pred. No. 2e-07;
Matches 98; Conservative 73; Mismatches 184; Indels 77; Gaps 20;

QY 110 LGNVLYLDRTYTAGVLLDQQHGVKDRGAGLLQSFVCSFVAAPIFGYLGDRFNRKV 169
Db 42 LGYVFDGDFMMIFYLHIKADILGIDTIDQATLIGTVAFIARPIGGGFFGAMADKYGRKP 101
QY 170 ILSGIFFWSAVTFSSSFIPOQYFELLVLSRGLVIGEASVSTIAPTIIIGDLFTKNTRTL 229
Db 102 MMWAIPIYSGTGLSIATNLY--MLAVCRFIVGLGMSGEYACASTYAVESWPKNLQSK 159
QY 230 MLSVYFPAIPLGSGLGVTGSSVKQAAGDWHALRVSPVLGMITGTLILILVPAKRGHA 289
Db 160 ASAPLVSGFSGVN--IIAAQIIPQFAEVYGM--RNSFFIGLLPVLVLWIRKSAPESQ- 213
QY 290 DQLGDQLKARTSWLRDMKALIRNSYVFSLSATSAVSF----ATGALGMWIPLYLHR--- 342
Db 214 EWIEDKYDKSTFLS-----VFRKPHLSISMIVFLVCFLFGANWPINGLLPSYLDNGV 268
QY 343 AQVVQKTAETCNPPCGAKDSLIFGAI--TCFTGFLGVVTGAGATRCWLKQTQADPLVC 400
Db 269 NTVVISLMTIAG-----LGLTGTIFFGVGDKIGV-----KKAF 304
QY 401 AVGMLGSAIFIC-LIFVAAK--SSIVGAYICIFVGETILLFSNWAITADI--LMYVVIPT- 455
Db 305 VVGLITSFIFLCPFFISVKNSSLIG--LCLF---GLMFTNLGIAGLVPKFIYDYPTKL 359
QY 456 RATAVALQSTSHLLDAGSPYLIIGFISDLIROSTKDSPLWEFLSLGYALMLCPFFVVVLG 515
Db 360 RGLGTGL-IYNLGATGGMAPVLATYISG-----YYGLGVSL----FIVTVA 401
QY 516 GMEFLATLFFV 527
Db 402 ---FSALLILLV 410
```

Search completed: October 23, 2004, 13:16:46  
Job time : 28 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 12:53:58 ; Search time 88 Seconds  
(without alignments)  
2082.357 Million cell updates/sec

Title: US-10-085-198-48

Perfect score: 2920

Sequence: 1 MMCLCASAAGGABEEAD.....VHQRGPGPTALAHRRVVGAS 566

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1364641 seqs, 323758627 residues

Total number of hits satisfying chosen parameters: 1364641

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09D\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2920	100.0	566	15 US-10-085-198-48	Sequence 48, Appl
2	1978	67.7	398	14 US-10-168-651-9	Sequence 9, Appl
3	1298	44.5	528	13 US-10-024-623-5	Sequence 5, Appl
4	1298	44.5	528	14 US-10-154-419-55	Sequence 55, Appl
5	1298	44.5	528	14 US-10-146-733-50	Sequence 50, Appl
6	1298	44.5	528	16 US-10-648-593-251	Sequence 251, Appl
7	1256.5	43.0	507	15 US-10-403-571-52	Sequence 52, Appl
8	1222	41.8	455	15 US-10-264-237-2641	Sequence 2641, Ap
9	1167	40.0	476	14 US-10-168-651-12	Sequence 12, Appl
10	1167	40.0	476	15 US-10-169-395-126	Sequence 126, Appl
11	1084	37.1	512	9 US-09-860-232A-2	Sequence 2, Appl
12	1084	37.1	512	14 US-10-297-022-23	Sequence 23, Appl
13	976	33.4	193	10 US-09-826-734-160	Sequence 160, Appl

14	855	29.3	531	14	US-10-154-419-81	Sequence 81, Appl
15	855	29.3	531	14	US-10-146-733-31	Sequence 31, Appl
16	779	26.7	299	10	US-09-809-391-396	Sequence 396, App
17	779	26.7	299	10	US-09-882-171-396	Sequence 396, App
18	779	26.7	299	14	US-10-164-861-396	Sequence 396, App
19	672	23.0	424	10	US-09-809-391-555	Sequence 555, App
20	672	23.0	424	10	US-09-882-171-555	Sequence 555, App
21	672	23.0	424	14	US-10-164-861-555	Sequence 555, App
22	644	22.1	329	15	US-10-108-260A-3290	Sequence 3290, Ap
23	559	19.1	343	14	US-10-097-065-395	Sequence 395, App
24	559	19.1	343	14	US-10-372-876-395	Sequence 395, App
25	553	18.9	304	14	US-10-097-065-153	Sequence 153, App
26	553	18.9	304	14	US-10-372-876-153	Sequence 153, App
27	534	18.3	106	10	US-09-826-734-246	Sequence 246, App
28	406	13.9	219	14	US-10-097-065-394	Sequence 394, App
29	406	13.9	219	14	US-10-372-876-394	Sequence 394, App
30	406	13.9	220	14	US-10-097-065-248	Sequence 248, App
31	406	13.9	220	14	US-10-372-876-248	Sequence 248, App
32	401	13.7	80	10	US-09-826-734-132	Sequence 132, App
33	396	13.6	136	10	US-09-809-391-554	Sequence 554, App
34	396	13.6	136	10	US-09-882-171-554	Sequence 554, App
35	396	13.6	136	14	US-10-164-861-554	Sequence 554, App
36	357.5	12.2	257	9	US-09-860-232A-19	Sequence 19, Appl
37	276	9.5	118	15	US-10-276-774-2582	Sequence 2582, Ap
38	268	9.2	412	15	US-10-425-114-49922	Sequence 49922, A
39	252	8.6	453	14	US-10-369-493-23684	Sequence 23684, A
40	252	8.6	480	16	US-10-437-963-125320	Sequence 125320,
41	221	7.6	243	15	US-10-424-599-195951	Sequence 195951,
42	219	7.5	546	14	US-10-369-493-21873	Sequence 21873, A
43	218.5	7.5	549	9	US-09-946-763-2	Sequence 2, Appl
44	218.5	7.5	549	9	US-09-738-626-3926	Sequence 3926, Ap
45	216.5	7.4	417	9	US-09-815-242-10592	Sequence 10592, A

#### ALIGNMENTS

#### RESULT 1

US-10-085-198-48  
; Sequence 48, Application US/10085198  
; Publication No. US20040009907A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsbrook et al.  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-279  
; CURRENT APPLICATION NUMBER: US/10/085,198  
; CURRENT FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 60/271,646  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 60/276,401  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/311,981  
; PRIOR FILING DATE: 2001-08-13  
; PRIOR APPLICATION NUMBER: 60/312,858  
; PRIOR FILING DATE: 2001-08-16  
; PRIOR APPLICATION NUMBER: 60/271,840  
; PRIOR FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: 60/277,324  
; PRIOR FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/286,096  
; PRIOR FILING DATE: 2001-04-21  
; PRIOR APPLICATION NUMBER: 60/299,695  
; PRIOR FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: 60/315,614  
; PRIOR FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/272,405  
; PRIOR FILING DATE: 2001-02-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 653  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 48  
; LENGTH: 566  
; TYPE: PRT

Sat Oct 23 17:45:11 2004

; ORGANISM: Homo sapiens  
US-10-085-198-48  
Query Match 100.0%; Score 2920; DB 15; Length 566;  
Best Local Similarity 100.0%; Pred. No. 8.4e-216;  
Matches 566; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNCLECAAAAGAEERERRRRGAQAGAGCGCCGARGAGAGAGVSAAGDEVQTL 60  
DB 1 MNCLECAAAAGAEERERRRRGAQAGAGCGCCGARGAGAGAGVSAAGDEVQTL 60  
QY 61 SSVRRAPPGCTGCTPCCATAGPQAQPKPASLGRGCAAAAILSLGNVNLRYDR 120  
DB 61 SSVRRAPPGCTGCTPCCATAGPQAQPKPASLGRGCAAAAILSLGNVNLRYDR 120  
QY 121 TVAGVLLDQQHFGVKDRGAGLLQSVFICSFVVAAPIFGYLGDNRNKRKVLSCGIFFWA 180  
DB 121 TVAGVLLDQQHFGVKDRGAGLLQSVFICSFVVAAPIFGYLGDNRNKRKVLSCGIFFWA 180  
QY 181 VTFSSSFIPOQYFWLLVLSRGLVIGEASYSIIAPTIIIGDLFTKTRTLMLSVFFFAIPL 240  
DB 181 VTFSSSFIPOQYFWLLVLSRGLVIGEASYSIIAPTIIIGDLFTKTRTLMLSVFFFAIPL 240  
QY 241 GSGLGVIITGSSVKQAAGDWHMALRVSPVLGMITGTLILVLPATKRGHADQLGDQKART 300  
DB 241 GSGLGVIITGSSVKQAAGDWHMALRVSPVLGMITGTLILVLPATKRGHADQLGDQKART 300  
QY 301 SWLRDMKALIRNSYVFFSSLSATSAVSFATGALGMWIPLYLHRAQVQVQKTAETCNSPPCGA 360  
DB 301 SWLRDMKALIRNSYVFFSSLSATSAVSFATGALGMWIPLYLHRAQVQVQKTAETCNSPPCGA 360  
QY 361 KDSLIIFGATCTFTGFLGVVTGAGATRCWLKTRQADPLVCVGMGLSAIFICLIFFVAKS 420  
DB 361 KDSLIIFGATCTFTGFLGVVTGAGATRCWLKTRQADPLVCVGMGLSAIFICLIFFVAKS 420  
QY 421 SIYGAVICIFVGTLLFSNWAITADILMVYVITRATAVALQSFTSHLLGDAGSPYLIG 480  
DB 421 SIYGAVICIFVGTLLFSNWAITADILMVYVITRATAVALQSFTSHLLGDAGSPYLIG 480  
QY 481 FISDLIRQSTKDSPLMEFLSLGYALMLCPVVLGMMFFLATALFFVSDRARAQHLGER 540  
DB 481 FISDLIRQSTKDSPLMEFLSLGYALMLCPVVLGMMFFLATALFFVSDRARAQHLGER 540  
QY 541 RAGVRVHHQRGPGGTALAHRVVGAS 566  
DB 541 RAGVRVHHQRGPGGTALAHRVVGAS 566

RESULT 2  
US-10-168-651-9  
; Sequence 9, Application US/10168651  
; Publication No. US20030171275A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: BURFORD, Neil  
; APPLICANT: AU-YOUNG, Janice  
; APPLICANT: LU, Dyung Aina M.  
; APPLICANT: YANG, Junming  
; APPLICANT: REDDY, Roopa  
; APPLICANT: LAL, Preeti  
; APPLICANT: HILLMAN, Jennifer L.  
; APPLICANT: AZIMZAI, Yalda  
; APPLICANT: YUE, Henry  
; APPLICANT: NGUYEN, Dannel B.  
; APPLICANT: YAO, Monique G.  
; APPLICANT: GANDHI, Ameen R.  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: KHAN, Farrah A.  
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS  
; FILE REFERENCE: PI-0005 PCT  
; CURRENT APPLICATION NUMBER: US/10/168,651  
; CURRENT FILING DATE: 2002-06-21

; PRIOR APPLICATION NUMBER: 60/172,000; 60/176,083; 60/177,332; 60/178,572; 60/179,758;  
60/181,625  
; PRIOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-21; 2000-01-28; 2000-02-02;  
2000-02-10  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PERL Program  
; SEQ ID NO 9  
; LENGTH: 398  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030171275A1 3232992CD1  
US-10-168-651-9  
Query Match 67.7%; Score 1978; DB 14; Length 398;  
Best Local Similarity 99.0%; Pred. No. 1.3e-143;  
Matches 384; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 152 MVAAPIFGYLGDNRNKRKVLSCGIFFWSAVTFSSSFIPOQYFWLLVLSRGLVIGEASYS 211  
DB 1 MVAAPIFGYLGDNRNKRKVLSCGIFFWSAVTFSSSFIPOQYFWLLVLSRGLVIGEASYS 60  
QY 212 TIAPTIIGDLFTKTRTLMLSVFFFAIPLGSGLGVIITGSSVKQAAGDWHMALRVSPVLGM 271  
DB 61 TIAPTIIGDLFTKTRTLMLSVFFFAIPLGSGLGVIITGSSVKQAAGDWHMALRVSPVLGM 120  
QY 272 ITGTLLILVLPATKRGHADQLGDQKARTSWLRDMKALIRNSYVFFSSLSATSAVSFATGA 331  
DB 131 ITGTLLILVLPATKRGHADQLGDQKARTSWLRDMKALIRNSYVFFSSLSATSAVSFATGA 180  
QY 332 LGMWIPLYLHRAQVQVQKTAETCNSPPCGAKDSLIFGATCTFTGFLGVVTGAGATRCWLK 391  
DB 181 LGMWIPLYLHRAQVQVQKTAETCNSPPCGAKDSLIFGATCTFTGFLGVVTGAGATRCWLK 240  
QY 392 TORADPLVCVGMGLSAIFICLIFFVAKSSIVGAYICIFVGTLLFSNWAITADILMVYV 451  
DB 241 TORADPLVCVGMGLSAIFICLIFFVAKSSIVGAYICIFVGTLLFSNWAITADILMVYV 300  
QY 452 IPTTRATAVALQSFTSHLLGDAGSPYLIGFISDLIRQSTKDSPLMEFLSLGYALMLCPPV 511  
DB 301 IPTTRATAVALQSFTSHLLGDAGSPYLIGFISDLIRQSTKDSPLMEFLSLGYALMLCPPV 360  
QY 512 VVLGMMFFLATALFFVSDRARAQHLGE 539  
DB 361 VVLGMMFFLATALFFVSDRARAQHLGE 388

RESULT 3  
US-10-024-623-5  
; Sequence 5, Application US/10024623  
; Publication No. US20020187524A1  
; GENERAL INFORMATION:  
; APPLICANT: Curtiss, Rory A. J.  
; TITLE OF INVENTION: 8099, 46455, 54414, 53763, 67076, 67102, 44181.  
; TITLE OF INVENTION: 67084PL, AND 67084 ALT, HUMAN PROTEINS AND METHODS OF  
; TITLE OF INVENTION: USE THEREOF  
; FILE REFERENCE: MNI-214CP  
; CURRENT APPLICATION NUMBER: US/10/024,623  
; CURRENT FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/256,240  
; PRIOR FILING DATE: 2000-12-15  
; PRIOR APPLICATION NUMBER: US 60/256,588  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: US 60/258,028  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 528  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-024-623-5



Query Match 44.5%; Score 1298; DB 13; Length 528;  
Best Local Similarity 53.4%; Pred. No. 3.1e-91;  
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;

QY 53 AGDEVQTLSSVRRAPTGP-PGTGTPGCAATAGKGAQPKPASLGR-----GRGA-A 104  
DB 2 AGSTAPFLSQADDDPGVPVGTGPGTGNPKSEBEVDPDQGLQRTGLSPGRSALI 61

QY 105 AAILSLGNVLNLYDRYTVAGVLLDIOQHFGVKDRAGLQSVFICSFMAAPIFGYLGDR 164  
DB 62 VAVLCYNLLNLYMDRFTVAGVLPDIEQFFNIGDSSGLIQTVFISSYMWLAPVFGYLGDR 121

QY 165 FNRKVIILSCGIFPWSAVTFSSSFTPOQYFWLLVLSRGLVGEASYSYTIAPTIIGDLFTK 224  
DB 122 YNRKYLCCGGIAFWLSVTLGSSFTPGSHFWLLLLTRGLVGVGEASYSYTIAPTLIADLFVA 181

QY 225 NTRTLMLSVFEFALPLGSGLYGTSSVKQAGDHWALRVSPLVGMITGTLLIILVLPAT 284  
DB 182 DQSRMLUSIFYFAIPVSGGLGYTAGSKVKQAGDHWALRVTGPGVAVVLLFLVVRP 241

QY 285 KRGHADQLGDLKAR-TSWLRDMKALIRNRSYFSSLSATSAVSPATGALGMWIPLYLHRA 343  
DB 242 PRGAVERHSDLPPLNPTSMWADLRALARNPSFVLSLGTAVAFVTVGSLAWAPAFLLRS 301

QY 344 QVQKTAETC-NSPPCGAKOSLIFGATCTGTFGLGVVTVGAGATRWCLKTQRADPLVCV 402  
DB 302 RVVLGETPPCLPGDSCSSDSLIFGLITCLTGLVGLGVGEISRRLRHNSPRADPLVCAT 361

QY 403 GMLGSAIFCICLIFVAAKSSIVGAYICIFVGETLLFSNWAITADILMYVVIPTTRATAVAL 462  
DB 362 GLGSAAPFLFLSLACARGSIVATYIFIFIGETLLSMNWAIVADILLVYVPTRRSTAEF 421

QY 463 QSFTSHLLGDAGSPYLIGFTSDLIROSTKDSPLMEFLSLGVALMCLCPVVVLGGMFFLAT 522  
DB 422 QIVLSHLLGDAGSPYLIGLISDRLRNWPSPFLSEFRALQFSLMCLCAFVGLGGAFLGT 481

QY 523 ALFPVSDRAEAOHLGE--RRAGVR-----VVHQRG 551  
DB 482 AIFTEADRRRAQLHVQGLLHEAGSTDDRIVVPQRG 516

RESULT 4  
US-10-154-419-55  
; Sequence 55, Application US/10154419  
; Publication No. US20030143675A1  
; GENERAL INFORMATION:  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Gluckman, Maria A.  
; TITLE OF INVENTION: NOVEL TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615,  
; TITLE OF INVENTION: HMDA-1, TWIK-9, alpha2delta-4, 54414, AND 53763 MOLECULES  
; TITLE OF INVENTION: AND USES THEREFOR  
; FILE REFERENCE: MNI-248  
; CURRENT APPLICATION NUMBER: US/10/154,419  
; CURRENT FILING DATE: 2002-05-22  
; PRIOR APPLICATION NUMBER: US 60/195,993  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: US 60/199,799  
; PRIOR FILING DATE: 2000-04-26  
; PRIOR APPLICATION NUMBER: US 60/233,537  
; PRIOR FILING DATE: 2000-09-19  
; PRIOR APPLICATION NUMBER: US 60/235,018  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US 60/235,059  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US 60/256,240  
; PRIOR FILING DATE: 2000-12-15  
; PRIOR APPLICATION NUMBER: US 60/256,588  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: US 60/258,028  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: US 09/796,720

Query Match 44.5%; Score 1298; DB 14; Length 528;  
Best Local Similarity 53.4%; Pred. No. 3.1e-91;  
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;

QY 53 AGDEVQTLSSVRRAPTGP-PGTGTPGCAATAGKGAQPKPASLGR-----GRGA-A 104  
DB 2 AGSDTAPFLSQADDDPGVPVGTGPGTGNPKSEBEVDPDQGLQRTGLSPGRSALI 61

QY 105 AAILSLGNVLNLYDRYTVAGVLLDIOQHFGVKDRAGLQSVFICSFMAAPIFGYLGDR 164  
DB 62 VAVLCYNLLNLYMDRFTVAGVLPDIEQFFNIGDSSGLIQTVFISSYMWLAPVFGYLGDR 121

QY 165 FNRKVIILSCGIFPWSAVTFSSSFTPOQYFWLLVLSRGLVGEASYSYTIAPTIIGDLFTK 224  
DB 122 YNRKYLCCGGIAFWLSVTLGSSFTPGSHFWLLLLTRGLVGVGEASYSYTIAPTLIADLFVA 181

QY 225 NTRTLMLSVFEFALPLGSGLYGTSSVKQAGDHWALRVSPLVGMITGTLLIILVLPAT 284  
DB 182 DQSRMLUSIFYFAIPVSGGLGYTAGSKVKQAGDHWALRVTGPGVAVVLLFLVVRP 241

QY 285 KRGHADQLGDLKAR-TSWLRDMKALIRNRSYFSSLSATSAVSPATGALGMWIPLYLHRA 343  
DB 242 PRGAVERHSDLPPLNPTSMWADLRALARNPSFVLSLGTAVAFVTVGSLAWAPAFLLRS 301

QY 344 QVQKTAETC-NSPPCGAKOSLIFGATCTGTFGLGVVTVGAGATRWCLKTQRADPLVCV 402  
DB 302 RVVLGETPPCLPGDSCSSDSLIFGLITCLTGLVGLGVGEISRRLRHNSPRADPLVCAT 361

QY 403 GMLGSAIFCICLIFVAAKSSIVGAYICIFVGETLLFSNWAITADILMYVVIPTTRATAVAL 462  
DB 362 GLGSAAPFLFLSLACARGSIVATYIFIFIGETLLSMNWAIVADILLVYVPTRRSTAEF 421

QY 463 QSFTSHLLGDAGSPYLIGFTSDLIROSTKDSPLMEFLSLGVALMCLCPVVVLGGMFFLAT 522  
DB 422 QIVLSHLLGDAGSPYLIGLISDRLRNWPSPFLSEFRALQFSLMCLCAFVGLGGAFLGT 481

QY 523 ALFPVSDRAEAOHLGE--RRAGVR-----VVHQRG 551  
DB 482 AIFTEADRRRAQLHVQGLLHEAGSTDDRIVVPQRG 516

RESULT 4  
US-10-154-419-55  
; Sequence 55, Application US/10154419  
; Publication No. US20030143675A1  
; GENERAL INFORMATION:  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Gluckman, Maria A.  
; APPLICANT: Meyers, Rachel E.  
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,  
; TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8099, 46455, 54414, 53763,  
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FBH58295FL, 57255,  
; TITLE OF INVENTION: AND 57255ALT MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-249  
; CURRENT APPLICATION NUMBER: US/10/154,419  
; CURRENT FILING DATE: 2002-05-22  
; PRIOR APPLICATION NUMBER: US 60/195,993  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: US 60/199,799  
; PRIOR FILING DATE: 2000-04-26  
; PRIOR APPLICATION NUMBER: US 60/233,537  
; PRIOR FILING DATE: 2000-09-19  
; PRIOR APPLICATION NUMBER: US 60/235,018  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US 60/235,059  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US 60/256,240  
; PRIOR FILING DATE: 2000-12-15  
; PRIOR APPLICATION NUMBER: US 60/256,588  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: US 60/258,028  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: US 09/796,720

```
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648.593
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 251
; LENGTH: 528
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-648-593-251

Query Match      44.5%; Score 1298; DB 16; Length 528;
Best Local Similarity 53.4%; Pred. No. 3.1e-91;
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;

QY 53 AGDEVQTLGSGVRRAPTGP-PGTPGPGCAATAGKGAQPKPASLGR-----GRGA-A 104
DB 2 AGSDTAPFLSQADDDPDGPGVPGTGLPGSTGNPKSEPEVPDQEGLQRIITGLSPGRSALI 61
QY 105 AAILSLGNVNLDRYTVAGVLLDIOQHFGVKDRGAGLLQSVFICSFMAAIPFGYLGDR 164
DB 62 VAVLCYINLNMDFRTVAGVLPDIEQFFNIGDSSSGLIQTFFISSYMLAPVFGYLGDR 121
QY 165 FNRKVLSCGIPFWSAVTFSSFIPOQYFWLLVLSRGLVGVGEASYSYTIPTIIGDLFTK 224
DB 122 YNRKYLKMGCGIAFWSLVTLGSSFIPEGHFHFWLLLTTRGLVGVGEASYSYTIPTLIADLFVA 181
QY 225 NTRTLMLSVFFPAIPGLSGLYITGSSVKQAAGDHWALRVSPVLGMITGTLLILVLPAT 284
DB 182 DQSRMLSIFFAIPVGSGLGYTAGSKVKDMAGDHWALRVTPGLGVVAVLLFLVVRP 241
QY 285 KRGHADQLGDLKAR-TSWLRDMKALIRNSYVFSLSLATSVAVSFATGALGMWIPLYLHRA 343
DB 242 PRGAVERHSDLPPLNPTSWADLRALARNPSFVLSLGTFAVAVFTGSLAWAPAFLLRS 301
QY 344 QVVQKTAETC-NSPPCGAKDSLIFGAITCTFTGLVGVVGTGATRWCLKTQRADPLVCV 402
DB 302 RVVLGETPPCLPGDSCSSDSLIFGLITCLTGLVGLGVGEISRRLRHNSPRADPLVCAT 361
QY 403 GMLGSAIFICLIFFVAKSSIVGAYICIFVGETLLFSNWAITADILMYVVIPTRRATAVAL 462
DB 362 GLLGSAPFLFSLACARGISVATYIFIFIGETLLSNWAIADILLYVVIPTRRSTAEAF 421
QY 463 QSFTSHLLGDAGSPYLIGFISDLIRQSTKDSPLWEFLSLGYALMLCPFFVVLGGMFFLAT 522
DB 422 QIVLSHLGLDAGSPYLIGLISDRLRNWPSPFLSEFRALQFSLMLCAFVGALGGAFLGT 481
QY 523 ALFFVSDRARAEOHLGE--RRAGVR-----VVHQRG 551
DB 482 AIFIEADRRRAQLHVQGLLHEAGSTDDRIWVPQRG 516

RESULT 7
US-10-403-571-52
; Sequence 52, Application US/10403571
; Publication No. US20040068763A1
; GENERAL INFORMATION:
; APPLICANT: Hopkins, Nancy
; APPLICANT: Gollings, Gregory
; APPLICANT: Amsterdam, Adam
; APPLICANT: Sun, Zhaoxia
; TITLE OF INVENTION: Developmental Mutations in Zebrafish
; FILE REFERENCE: 01997/539002
; CURRENT APPLICATION NUMBER: US/10/403,571
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US 60/368,760
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 507
```

```
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/828,035
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 09/833,081
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 09/843,128
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 09/957,683
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 09/964,252
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/964,256
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 10/024,623
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 528
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-733-50

Query Match      44.5%; Score 1298; DB 14; Length 528;
Best Local Similarity 53.4%; Pred. No. 3.1e-91;
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;

QY 53 AGDEVQTLGSGVRRAPTGP-PGTPGPGCAATAGKGAQPKPASLGR-----GRGA-A 104
DB 2 AGSDTAPFLSQADDDPDGPGVPGTGLPGSTGNPKSEPEVPDQEGLQRIITGLSPGRSALI 61
QY 105 AAILSLGNVNLDRYTVAGVLLDIOQHFGVKDRGAGLLQSVFICSFMAAIPFGYLGDR 164
DB 62 VAVLCYINLNMDFRTVAGVLPDIEQFFNIGDSSSGLIQTFFISSYMLAPVFGYLGDR 121
QY 165 FNRKVLSCGIPFWSAVTFSSFIPOQYFWLLVLSRGLVGVGEASYSYTIPTIIGDLFTK 224
DB 122 YNRKYLKMGCGIAFWSLVTLGSSFIPEGHFHFWLLLTTRGLVGVGEASYSYTIPTLIADLFVA 181
QY 225 NTRTLMLSVFFPAIPGLSGLYITGSSVKQAAGDHWALRVSPVLGMITGTLLILVLPAT 284
DB 182 DQSRMLSIFFAIPVGSGLGYTAGSKVKDMAGDHWALRVTPGLGVVAVLLFLVVRP 241
QY 285 KRGHADQLGDLKAR-TSWLRDMKALIRNSYVFSLSLATSVAVSFATGALGMWIPLYLHRA 343
DB 242 PRGAVERHSDLPPLNPTSWADLRALARNPSFVLSLGTFAVAVFTGSLAWAPAFLLRS 301
QY 344 QVVQKTAETC-NSPPCGAKDSLIFGAITCTFTGLVGVVGTGATRWCLKTQRADPLVCV 402
DB 302 RVVLGETPPCLPGDSCSSDSLIFGLITCLTGLVGLGVGEISRRLRHNSPRADPLVCAT 361
QY 403 GMLGSAIFICLIFFVAKSSIVGAYICIFVGETLLFSNWAITADILMYVVIPTRRATAVAL 462
DB 362 GLLGSAPFLFSLACARGISVATYIFIFIGETLLSNWAIADILLYVVIPTRRSTAEAF 421
QY 463 QSFTSHLLGDAGSPYLIGFISDLIRQSTKDSPLWEFLSLGYALMLCPFFVVLGGMFFLAT 522
DB 422 QIVLSHLGLDAGSPYLIGLISDRLRNWPSPFLSEFRALQFSLMLCAFVGALGGAFLGT 481
QY 523 ALFFVSDRARAEOHLGE--RRAGVR-----VVHQRG 551
DB 482 AIFIEADRRRAQLHVQGLLHEAGSTDDRIWVPQRG 516

RESULT 6
US-10-648-593-251
; Sequence 251, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
```

[illegible]

```
Db      2 AGSDTAPFLSQADDDPDDGVPVPGTGLPGSTGNPKSBEPEVPDQEGLQRIITGLSPQRSALI 61
Qy      105 AAILSGVNLVLDRTYVAGVLLDIOQHFGVKDRGAGLLQSVFICSEFWAAPIFGYLGDR 164
Db      62 VAVLCYINLLNMDRFTVAGVLPDIEQFFNIGDSSGLIQTFFISSYMWLAPVFGYLGDR 121
Qy      165 FNRKVLSCGIFPWSAVTFSSSFIPOQYFWLLVLSRGLVIGGEASVSTIAPTIIIGDLFTK 224
Db      122 YNRKYLCCGIAFWSLVTLGSSFIPEGEHFWLLLLTRGLVGVGEASVSTIAPTIIADLFVA 181
Qy      225 NTRTLMLSVFFFAIPGLSGGLYITGSSVKQAAGDWHWALRVSPVLGMITGTILILVPA 284
Db      182 DQSRMLSFIFYFAIPVGSGLYIAGSKVDMAGDWHWALRVTPGLGVAVLLLELVVREP 241
Qy      285 KRGHADQLGDQKAR-TSWLRDMKALIRNRSVVFSSLSATSVSFATGALGMWIPLYLHRA 343
Db      242 PRGAVERHSDLPPLNPTSWADLRARN----- 270
Qy      344 QVVQKTAETCNSPPCGAKDSLIFGAITCFTGFLGVVTVGAGATRWCLKTQRADPLVCVG 403
Db      271 -----LIFGLITCLTGVGLGVGLVEISRRLRHNSPRADPLVCATG 310
Qy      404 MLGSAIFICLIFVAAKSSIVGAYICIFVGETLLFSNWAITADILMVVVIPTRRATAVALQ 463
Db      311 LLGSAPFLFLSLACARGSVATYIFIFIGETLLSNMWAIVADILLYVVIPTRRSTAEAFQ 370
Qy      464 SFTSHLLGDAGSPYLIGFISDLIRQSTKDSPLWEFLSLGYALMCLPCFVVVLGGMFLATA 523
Db      371 IVLSHLLGDAGSPYLIGLISDLRNRNWPSPFLSEFRALQFSLMCLCAFGALGAAFLGTA 430
Qy      524 LFFVSDRARAEOHLGE--RRAGVR-----VVHQRG 551
Db      431 IFIEADRRRAQLHVQGLLHEAGSTDDRIVVVPRG 464
```

```
RESULT 10
US-10-169-395-126
; Sequence 126, Application US/10169395
; Publication No. US20040034192A1
; GENERAL INFORMATION:
; APPLICANT: KATO, Seiichi
; APPLICANT: KIMURA, Tomoko
; TITLE OF INVENTION: HUMAN PROTEINS HAVING HYDROPHOBIC DOMAINS AND DNAs ENCODING
; FILE REFERENCE: 01997.015100.US
; CURRENT APPLICATION NUMBER: US/10/169,395
; PRIOR FILING DATE: 2002-11-29
; PRIOR APPLICATION NUMBER: JP 2000-585
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: JP 2000-588
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: JP 2000-2299
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-26862
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: JP 2000-58367
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/JP00/09359
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 150
; SEQ ID NO 126
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-169-395-126

Query Match
Best Local Similarity 40.0%; Score 1167; DB 15; Length 476;
Matches 254; Conservative 53; Mismatches 141; Indels 66; Gaps 7;

Qy      53 AGDEVOTLSGSVRRAPTGP-PGTPTPGCAATAKAGGAAQPKPASLGR-----GRGA-A 104
Db      2 AGSDTAPFLSQADDDPDDGVPVPGTGLPGSTGNPKSBEPEVPDQEGLQRIITGLSPQRSALI 61
```

```
Qy      105 AAILSGVNLVLDRTYVAGVLLDIOQHFGVKDRGAGLLQSVFICSEFWAAPIFGYLGDR 164
Db      62 VAVLCYINLLNMDRFTVAGVLPDIEQFFNIGDSSGLIQTFFISSYMWLAPVFGYLGDR 121
Qy      165 FNRKVLSCGIFPWSAVTFSSSFIPOQYFWLLVLSRGLVIGGEASVSTIAPTIIIGDLFTK 224
Db      122 YNRKYLCCGIAFWSLVTLGSSFIPEGEHFWLLLLTRGLVGVGEASVSTIAPTIIADLFVA 181
Qy      225 NTRTLMLSVFFFAIPGLSGGLYITGSSVKQAAGDWHWALRVSPVLGMITGTILILVPA 284
Db      182 DQSRMLSFIFYFAIPVGSGLYIAGSKVDMAGDWHWALRVTPGLGVAVLLLELVVREP 241
Qy      285 KRGHADQLGDQKAR-TSWLRDMKALIRNRSVVFSSLSATSVSFATGALGMWIPLYLHRA 343
Db      242 PRGAVERHSDLPPLNPTSWADLRARN----- 270
Qy      344 QVVQKTAETCNSPPCGAKDSLIFGAITCFTGFLGVVTVGAGATRWCLKTQRADPLVCVG 403
Db      271 -----LIFGLITCLTGVGLGVGLVEISRRLRHNSPRADPLVCATG 310
Qy      404 MLGSAIFICLIFVAAKSSIVGAYICIFVGETLLFSNWAITADILMVVVIPTRRATAVALQ 463
Db      311 LLGSAPFLFLSLACARGSVATYIFIFIGETLLSNMWAIVADILLYVVIPTRRSTAEAFQ 370
Qy      464 SFTSHLLGDAGSPYLIGFISDLIRQSTKDSPLWEFLSLGYALMCLPCFVVVLGGMFLATA 523
Db      371 IVLSHLLGDAGSPYLIGLISDLRNRNWPSPFLSEFRALQFSLMCLCAFGALGAAFLGTA 430
Qy      524 LFFVSDRARAEOHLGE--RRAGVR-----VVHQRG 551
Db      431 IFIEADRRRAQLHVQGLLHEAGSTDDRIVVVPRG 464

RESULT 11
US-09-860-232A-2
; Sequence 2, Application US/09860232A
; Patent No. US20020028494A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: 57256 AND 58289, NOVEL HUMAN
; FILE REFERENCE: 381552001500
; CURRENT APPLICATION NUMBER: US/09/860,232A
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/205,288
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-860-232A-2

Query Match
Best Local Similarity 37.1%; Score 1084; DB 9; Length 512;
Matches 223; Conservative 75; Mismatches 163; Indels 8; Gaps 3;

Qy      75 PGTPTGCAATAKAGGAAQPKPA-----SLGRGRG-AAAAISLGNVNLVLDRTYVAGVLLD 128
Db      12 FPGGGLQGQSPGPGRCPPPTPTPSWSLPPWRAVYVAAAALVCYNLLNMYNMFIIAGVLLD 71
Qy      129 IQQHFGVKDRGAGLLQSVFICSEFWAAPIFGYLGDRFNRKVLSCGIFPWSAVTFSSSFI 188
Db      72 IQEVFQISDNHAGLLQTVFVSCLLSAFVFGVLSRHSRKATMSFGIILLWSGAGSSSFI 131
Qy      189 PQQYFWLLVLSRGLVIGGEASVSTIAPTIIIGDLFTKTRTLMLSVFFFAIPGLSGGLYIT 248
Db      132 SPTKSWLFFLSRGIVGTGSASVSTIAPTIVGLDFVRDQTRVLAIFYIFIVGSGGLGYVL 191
Qy      249 GSSVQKAAAGDWHWALRVSPVLGMITGTILILVPAITKRGHADQLGAKA--RTSWLRDM 306
```

Db 192 GSAVTMLTGNWRWALRVMPCLAEVALILLILLVDPDPGRGAETGEGAVGGRSSWCEDV 251  
Qy 307 KALIRNRSYVFPSSLSATSAVSFATGALGMWIPLYLHRAQVQVQKTAETCNPSPPCGAKDSLIF 366  
Db 252 RYLGNKWSFVMSLTVTAMAFVTGALGFWAPKFLLEARVHGLQPPCFQPCSNPDSLIF 311  
Qy 367 GAITCTFTGFLGVVGTGAGATRCWLKTRQADPLCAVGMGLSAIFICLIFVAAKSSIVGAY 426  
Db 312 GALTMTGTVIGVILGAESRRYKVIIPGAEPLICASSLLATAPCLYLALVLAPTTLASY 371  
Qy 427 ICIFVGETLLFSSNWAITADILMVVITPRRATAVALQSFTSHLIGDAGSPYLIGFISDLI 486  
Db 372 VFGLGELLSSCNWVAVDILLVVVPRCRGTAEALQITVGHILGDAGSPYLIGLSSVL 431  
Qy 487 RQSTKDSPLWELSLGVALMCLPFFVVLGGMFFLATALFFVSDRAREQ 535  
Db 432 RARRPDSYLQFRSLQOSFLCCAFVIALGGCFLLTALYLERDETRAWQ 480

## RESULT 12

US-10-297-022-23

; Sequence 23, Application US/10297022

; Publication No. US20030216310A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.

; APPLICANT: THORNTON, Michael

; APPLICANT: WALIA, Narinder K.

; APPLICANT: YUE, Henry

; APPLICANT: NGUYEN, Dannel B.

; APPLICANT: LAL, Preeti

; APPLICANT: GANDHI, Aneena R.

; APPLICANT: TRIBOLEY, Catherine M.

; APPLICANT: YAO, Monique G.

; APPLICANT: RAMKUMAR, Javalaxmi

; APPLICANT: AU-YOUNG, Janice

; APPLICANT: LU, Yan

; APPLICANT: TANG, Y. Tom

; APPLICANT: AZIMZAI, Yalda

; APPLICANT: BRUNS, Christopher M.

; APPLICANT: GRIFFIN, Jennifer A.

; APPLICANT: YANG, Junming

; APPLICANT: BAUGHN, Mariah R.

; APPLICANT: SANJANWALA, Madhu S.

; APPLICANT: RAUMANN, Brigitte E.

; APPLICANT: LEE, Ernestine A.

; APPLICANT: HAFALIA, April

; APPLICANT: GREENE, Barrie D.

; APPLICANT: KHAN, Farrah A.

; APPLICANT: KEARNEY, Liam

; APPLICANT: ELLIOTT, Vicky S.

; APPLICANT: SEILHAMER, Jeffrey J.

; APPLICANT: POLICKY, Jennifer L.

; APPLICANT: BOROWSKY, Mark L.

; APPLICANT: BURFORD, Neil

; APPLICANT: DING, Li

; APPLICANT: LU, Dyang Aina M.

; APPLICANT: HILLMAN, Jennifer L.

; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS

; FILE REFERENCE: FI-0109 PCT

; CURRENT APPLICATION NUMBER: US/10/297,022

; CURRENT FILING DATE: 2002-11-25

; PRIOR APPLICATION NUMBER: 60/208,424; 60/209,001; 60/210,588; 60/212,335; 60/213,747

; PRIOR FILING DATE: 2000-05-26; 2000-06-01; 2000-06-08; 2000-06-16; 2000-06-22; 2000-06-24

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: PERL Program

; SEQ ID NO 23

; LENGTH: 512

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. US20030216310A1 1713377CD1

US-10-297-022-23

Query Match 37.1%; Score 1084; DB 14; Length 512;  
Best Local Similarity 47.5%; Pred. No. 8.3e-75;  
Matches 223; Conservative 75; Mismatches 163; Indels 8; Gaps 3;  
Qy 75 PGTGCGCATAKGCAQPKPA-----SLGRGRG-AAAAIISLGNVLNLYDRYTVAGVLLD 128  
Db 12 PGGGGLQGGQPGFGRCQPPITPTSTSLPPWRAVAAAVLCYNLLNMMFIIAGVLLD 71  
Qy 129 IQOHFGVKDRGAGLQSVFICSPFWAAAPIFGYLGDRNRKRVILSCGFFPFSVAVTFSSFI 188  
Db 72 IQEVFQISDNHAGLLQTVFVSCILLASAPVFGYLGDRHSRKATMSFGILLMSGAGLSSFI 131  
Qy 189 PQDYFLLVLSRGLVGIGEASYSYTIPTIIGDLFTKNTRTLMLSVFYFAIPLGSGLYIT 248  
Db 132 SPRYSMLFSLRGIVGTGSASYSYTIPTVLGDLFVRDQRTVLAVFYFIPVGSGLGYVL 191  
Qy 249 GSSVKQAGDWHWALRVSPVLGMITGTLIIILIPATKRGHADQLGDOLKA--RTSWLRDM 306  
Db 192 GSAVTMLTGNWRWALRVMPCLAEVALILLILLVDPDPGRGAETGEGAVGGRSSWCEDV 251  
Qy 307 KALIRNRSYVFPSSLSATSAVSFATGALGMWIPLYLHRAQVQVQKTAETCNPSPPCGAKDSLIF 366  
Db 252 RYLGNKWSFVMSLTVTAMAFVTGALGFWAPKFLLEARVHGLQPPCFQPCSNPDSLIF 311  
Qy 367 GAITCTFTGFLGVVGTGAGATRCWLKTRQADPLCAVGMGLSAIFICLIFVAAKSSIVGAY 426  
Db 312 GALTMTGTVIGVILGAESRRYKVIIPGAEPLICASSLLATAPCLYLALVLAPTTLASY 371  
Qy 427 ICIFVGETLLFSSNWAITADILMVVITPRRATAVALQSFTSHLIGDAGSPYLIGFISDLI 486  
Db 372 VFGLGELLSSCNWVAVDILLVVVPRCRGTAEALQITVGHILGDAGSPYLIGLSSVL 431  
Qy 487 RQSTKDSPLWELSLGVALMCLPFFVVLGGMFFLATALFFVSDRAREQ 535  
Db 432 RARRPDSYLQFRSLQOSFLCCAFVIALGGCFLLTALYLERDETRAWQ 480

## RESULT 13

US-09-826-734-160

; Sequence 160, Application US/09826734

; Publication No. US20030017457A1

; GENERAL INFORMATION:

; APPLICANT: Fernandes, Elma R.

; APPLICANT: Vernet, Corine A.M.

; APPLICANT: Mishra, Vishnu S.

; APPLICANT: Leach, Martin D.

; APPLICANT: Shimkets, Richard A.

; APPLICANT: Zerhusen, Bryan D.

; APPLICANT: Kekuda, Ramesha

; TITLE OF INVENTION: Novel Polynucleotides and Polypeptides Encoded Thereby

; FILE REFERENCE: 15966-754

; CURRENT APPLICATION NUMBER: US/09/826,734

; CURRENT FILING DATE: 2001-04-05

; PRIOR APPLICATION NUMBER: 60/195,576

; PRIOR FILING DATE: 2000-04-06

; NUMBER OF SEQ ID NOS: 270

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 160

; LENGTH: 193

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-826-734-160

Query Match 33.4%; Score 976; DB 10; Length 193;  
Best Local Similarity 98.4%; Pred. No. 5.2e-67;  
Matches 190; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 143 LQSVFICSPFWAAPIFGYLGDRNRKRVILSCGFFPFSVAVTFSSFPQQYFWLLVLSRGL 202

Db 1 LPAVFCSPFWAAPIFGYLGDRNRKRVILSCGFFPFSVAVTFSSFPQQYFWLLVLSRGL 60

Qy 203 VGIGEASYSYTIPTIIGDLFTKNTRTLMLSVFYFAIPLGSGLYITGSSVKQAAGDWHWA 262

Db 61 VGIGESYSTIAPTIIIGDLFTKNTRLMLSVFYFAIPGLGSLGYITGSSVKQAAGDWHWA 120  
QY 263 LRVSPVLGMITGTLILIVPATKRGHADQDGLKARTSWLRDMKALINRYSYVSSLAT 322  
Db 121 LRVSPVLGMITGTLILIVPATKRGHADQDGLKARPSWLRDMKALINRYSYVSSLAT 180  
QY 323 SAVSFATGALGMW 335  
Db 181 SAVSFATGALGMW 193  
RESULT 14  
US-10-154-419-81  
; Sequence 81, Application US/10154419  
; Publication No. US20030143675A1  
; GENERAL INFORMATION:  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Gluckman, Maria Alexandra  
; APPLICANT: Meyers, Rachel E.  
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,  
; TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8099, 46455, 54414, 53763,  
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FBH58295FL, 57255,  
; TITLE OF INVENTION: AND 57255salt MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-249  
; CURRENT APPLICATION NUMBER: US/10/154,419  
; CURRENT FILING DATE: 2002-05-22  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 99  
; SOFTWARE: FastSeq Version 4.0  
; SEQ ID NO 81  
; LENGTH: 531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-154-419-81  
Query Match 29.3%; Score 855; DB 14; Length 531;  
Best Local Similarity 40.0%; Pred. No. 3.4e-57;  
Matches 194; Conservative 82; Mismatches 173; Indels 36; Gaps 9;  
QY 71 PPGTGTPT-----GCAATAGGGAQOPKPSLGRGRGAAAAIILSGNVLYLDRTY 121  
Db 23 PPYTTTDPEDKIRNSNTATTASQPEFQ-----GCWTIVVAILFIINLLNMDRYT 75  
QY 122 VAGVLDIQHFGVKDGRAGLQSVFICSMVAAPIFGVLGRFNRKVLSCGIFWSAV 181  
Db 76 IAGVLNDVQTYNISDAWAGLIQTTPMVFIIIFSPICGFLGRYNNRKNWIFVVGIAIWVSA 135  
QY 182 TFSSSPIQOYFWLLVLSRGLVGIGESYSTIAPTIIIGDLFTKNTRLMLSVFYFAIPGLG 241  
Db 136 VFASFTIPSNQFWLFLFRGIVGIGESYALISPTVIADMTGVLRSRLMWFYFAIPFG 195  
QY 242 SGLGYITGSSVQAAAGDHWALRVSPVLGMITGTLILIVPATKRGHADQDGLKA--- 298  
Db 196 CGLGFFVVGSAVSWTGHWQWGRVVTGVLGIVCLLIIVFVREPERGKABERKEGTAASPE 255  
QY 299 RTSWLRDMKALINRYSYVSSLATSAVSFATGALGMWIFLYLHRAQVVKTAETCNSPPC 358  
Db 256 ATSYLDMDKLLSNATYVTSLSGYTATVFMVGTFLAWNPITIQYADSAARNGTITTEDQK- 314  
QY 359 GAKDSLIFGAITCFTGFLGVVWGA-CATRWCR-----LKTORADPLVCVAGMLGSAIF 410  
Db 315 -ANINLVFGLNLCVGGVLGVALGTLVSNMWSRGVGPFKHIQIVRADALVCAI---GAALC 370  
QY 411 ICLIFVA---AKSSIVGAYICIFVGETLLFSNWAITADILMVVPIPTRTATVALQSFTS 467  
Db 371 IPTLILAIQIESNNMNFAMGMLFICIVASSFNWATNVDLLSVVFPQRSSASSWQILIS 430  
QY 468 HLLGAGSPYLIGFTSLLRQSTKPSLWFEFLSLGVALMLCPVVVVLGCMFFLATALFPV 527  
Db 431 HMFGDAGSPYLIGLISDAIR-GNEDTAQAHYKSLVTSFWLCVGTLLVLSILFGISAITVV 489  
QY 528 SDRAR 532

Db 490 KDKAR 494  
RESULT 15  
US-10-146-733-31  
; Sequence 31, Application US/10146733  
; Publication No. US20030165891A1  
; GENERAL INFORMATION:  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Glucksmann, Maria A.  
; APPLICANT: Silos-Santiago, Inmaculada  
; TITLE OF INVENTION: NOVEL TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615,  
; TITLE OF INVENTION: HNMDA-1, TWIK-9, alpha2delta-4, 54414, AND 53763 MOLECULES  
; TITLE OF INVENTION: AND USES THEREFOR  
; FILE REFERENCE: MNI-248  
; CURRENT APPLICATION NUMBER: US/10/146,733  
; CURRENT FILING DATE: 2002-05-15  
; PRIOR APPLICATION NUMBER: US 60/185,938  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 09/515,520  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 09/518,866  
; PRIOR FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: US 60/195,734  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: US 60/195,993  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: US 60/199,799  
; PRIOR FILING DATE: 2000-04-26  
; PRIOR APPLICATION NUMBER: US 60/233,537  
; PRIOR FILING DATE: 2000-09-19  
; PRIOR APPLICATION NUMBER: US 60/235,018  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US 60/235,059  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US 60/256,240  
; PRIOR FILING DATE: 2000-12-15  
; PRIOR APPLICATION NUMBER: US 60/256,588  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: US 60/258,028  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: US 09/796,720  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 09/828,035  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: US 09/833,081  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: US 09/843,128  
; PRIOR FILING DATE: 2001-04-25  
; PRIOR APPLICATION NUMBER: US 09/957,683  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: US 09/964,252  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 09/964,256  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 10/024,623  
; PRIOR FILING DATE: 2001-12-17  
; NUMBER OF SEQ ID NOS: 85  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 31  
; LENGTH: 531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-146-733-31  
Query Match 29.3%; Score 855; DB 14; Length 531;  
Best Local Similarity 40.0%; Pred. No. 3.4e-57;  
Matches 194; Conservative 82; Mismatches 173; Indels 36; Gaps 9;  
QY 71 PPGTGTPT-----GCAATAGGGAQOPKPSLGRGRGAAAAIILSGNVLYLDRTY 121  
Db 23 PPYTTTDPEDKIRNSNTATTASQPEFQ-----GCWTIVVAILFIINLLNMDRYT 75



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 12:50:48 ; Search time 27 Seconds  
(without alignments)  
2016.988 Million cell updates/sec

Title: US-10-085-198-48  
Perfect score: 2920  
Sequence: 1 MMCLCASAAGGAEBEAD.....VHQRGPGGTALAHVRVVGAS 566  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : FIR\_79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	855	29.3	531	2	T19232	hypothetical prote
2	815	27.9	488	2	T19854	hypothetical prote
3	681	23.3	411	2	T20623	hypothetical prote
4	392.5	13.4	507	2	B84616	hypothetical prote
5	354.5	12.1	746	2	T05899	hypothetical prote
6	285	9.8	438	2	G87290	major facilitator
7	278	9.5	457	2	R3064	probable MFS trans
8	276	9.5	519	2	E87557	major facilitator
9	252	8.6	453	1	P62250	hypothetical 49.4
10	252	8.6	453	2	D91293	probable transport
11	252	8.6	453	2	G86134	hypothetical prote
12	247	8.5	444	2	D87557	major facilitator
13	219	7.5	546	2	S48313	hexose transport p
14	217	7.4	479	2	P87474	major facilitator
15	214	7.3	422	1	A69853	hexuronate transpo
16	205	7.0	541	1	MMBYH2	glucose transport
17	204.5	7.0	436	2	AG0272	probable integral
18	203.5	7.0	457	2	D71144	hypothetical prote
19	203	7.0	422	2	D72302	hypothetical prote
20	200	6.8	456	1	S27616	probable glucarate
21	196	6.7	472	2	B99372	multidrug resistanc
22	195.5	6.7	422	2	H69839	multidrug resistanc
23	195	6.7	490	2	E96810	probable transport
24	194	6.6	455	1	H69752	probable glucarat
25	194	6.6	458	1	YTESY8	tetracycline resis
26	194	6.6	477	2	B75409	multidrug-efflux t
27	192	6.6	472	2	F90355	multidrug resistanc
28	183.5	6.3	418	2	S76669	hypothetical prote
29	183.5	6.3	459	2	S42238	tetracyclin resist

30 183.5 6.3 521 2 T50023  
31 182.5 6.2 537 2 D45634  
32 182 6.2 372 2 D75094  
33 182 6.2 412 2 S39734  
34 182 6.2 444 2 R85974  
35 182 6.2 444 2 E91129  
36 182 6.2 444 2 C65102  
37 182 6.2 450 2 A65061  
38 181.5 6.2 586 2 B87408  
39 181 6.2 433 2 G86812  
40 181 6.2 452 2 A80861  
41 181 6.2 458 1 YTESRT  
42 180.5 6.2 458 2 S23743  
43 180.5 6.2 567 2 S31294  
44 179.5 6.1 372 2 H71070  
45 178 6.1 421 2 AD3393

hypothetical prote  
hypothetical prote  
transport protei  
chloramphenicol re  
probable galactara  
probable galactara  
probable glucarate  
tetracycline resis  
D-xylose proton-ey  
probable glucarate  
tetracycline resis  
hexose transport p  
hypothetical prote  
bicyclomycin resis

RESULT 1  
T19232  
hypothetical protein C13C4.5 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T19232  
R;Harris, B.  
submitted to the EMBL Data Library, March 1997  
A;Reference number: Z19095  
A;Accession: T19232  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-531 <WIL>  
A;Cross-references: UNIPROT:O01927; EMBL:Z92825; PIDN:CAB07311.1; GSPDB:GN00023; CESP:  
C;Genetics:  
A;Gene: CESP:C13C4.5  
A;Map position: 5  
A;Introns: 78/1; 145/3; 269/2; 341/3; 389/3; 411/3; 500/1

ALIGNMENTS

Query Match 29.3%; Score 855; DB 2; Length 531;  
Best Local Similarity 40.0%; Pred. No. 1.5e-50;  
Matches 194; Conservative 82; Mismatches 173; Indels 36; Gaps 9;

Qy 71 PPGTGP-----GCAATAGCGAQQPKPASLGRGAAAIISLGNVLNLDYRT 121  
Db 23 PPTPTDSDPKIRSNSTATTASQPEFQ-----GCWTIVVVAIIIFINLLNMDRYT 75  
Qy 122 VAGVLLDIQQHFGVKRGAGLQSVFICSFVMAAPIFGYLGDRFNKRVILSCGIPFFWSAV 181  
Db 76 IAGVLNDVQYNNISDAWAGLIQTFMVFFIIFSPICGFLGDRYRKNWIFVVGIAIWISA 135  
Qy 182 TFSSSFTPOQYFWLLVLSRGLVGIGEASYSYTIPTIGDLFTKNTRLMLSVFPAIPLG 241  
Db 136 VFASTPIPSNQFWLFLFRGIVGIGEASYSYTIPTIADMTFTGLSRMLMVFVFAIPFG 195  
Qy 242 SLGIYITGSSVKQAGDWHALRVSPVLGMITGTLILLILYPATKRGHADQGDQLKA--- 298  
Db 196 CGLGFVVGSASVASTGHWQMGVRVTGLVIGVCLLLIIVFVREPERGKABREKGAIASTE 255  
Qy 299 RTSLWRDMKALIRNRSVFFSSLSATSVSFATGALGMWIPLYLHRAQVQVQVQTAETCNSPPC 358  
Db 256 ATSVLDMDKLLSNATVTSLSGYTATVFWVGTLAWAPITIQVADSARRNGITEDOK- 314  
Qy 359 GAKDSLIFGATCTGFLGVVGTGA-GATRCR-----LKTORADPLVCVAVGLMGSIAIF 410  
Db 315 -ANINLVEGALTCVGVGLVGAIGTFLVSNMNSRGVGPFKHIQTVRADALVCAI---GAAC 370  
Qy 411 ICLIFVA---AKSIVGAYICIFVGETLLPSNWAITADILMYVVIPTRTAVALQSFST 467  
Db 371 IPTLILAIONIESNMNFAWGLMFCIVASSFNWATNDLLSVVVPQRRSSASSWQILIS 430



Db 10 PVTETTSRCYSTSSSTPLABLETVRSLEIVVESSSLSPVWLLVIFCIINLLNMDRGAI 69  
Qy 123 AG-----VLLDIQQHFGVDRGAGLIQSIFGSMVAAPFICGLGDRF 165  
Db 70 ASNGVNGSTRSCNDKCKTATGQGHENLSPEDGVLSSEFMVGLLIASIFASLAKR- 128  
Qy 166 NRKVLSCGIFPWSAVTF--SSSFIPOQYFWLLVLSRGLVGIGBASVSTIAPTIGDLFT 223  
Db 129 ----LIGVGLTVWTIAVLGCGSSPA---FWFVLCRMFVGVGSEAFISLAAPFIDNAP 180  
Qy 224 KNTETLMSVFPFALPLSGSLGYITGSSVKQ-----AAGDWHALRVSP--VLGMITGLI 277  
Db 181 QEQKAAMGLFIMCIPSGVALGYVGGVGHFNRVAFGEAVLMAPFAVLGEMKPLQ 240  
Qy 278 L-----ILVPATKRGHADQLGDLKARTSWLRDMKALIRNR 313  
Db 241 LKGSSETLKNNNLQVDNEIHDQFEVSIETSKSYANAV---FKSFTGFADKMVLYKEK 297  
Qy 314 SYFPSSLSATSAVSPATGALGMWIP-----LYLHRAQVQKTAETCNSPPCGAKDSLIFGA 368  
Db 298 VFVNVVLGYVSNFVIGAYSYWGPKAGYNIY-----KMKNAD-----MIFGA 339  
Qy 369 ITCTGFLGVVGTGAGATWCKLTQADPLVCAVGMGSAIFCLIFVAAKSSIVGAYIC 428  
Db 340 VTIICGVITLSGGFILDRTVATIPNAFKLLSGATFLG-AVF-CFTAFTLKS--LYGFIA 395  
Qy 429 IF-VGETLLFSNMAITADILMYVVIPTRRATAVALQSFTHSLGDAGSPYLIGFISDLIR 487  
Db 396 LFALGELLVATQAPVNVVCHCVKPSLRPLSMAISVAIHFGDVSPSPVLGVVQDHIN 455  
Qy 488 QSTKDSPLWFLSLGYALMLCPFVVVLGGMFFLATALFFVS-----DRARAEQ 535  
Db 456 SWRKTIT-----LILTSILFLAAIWFIGIFINSVDREFNOE 491

RESULT 5  
T05899  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 09-Apr-1999 #sequence\_revision 09-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T05899  
R:Bevan, M.; Brandt, P.; Dose, S.; Jarke, D.; Scharfe, M.; Schon, O.; Bancroft, I.; Mew  
submitted to the Protein Sequence Database, April 1998  
A:Reference number: Z15456  
A:Accession: T05899  
A:Molecule type: DNA  
A:Residues: 1-746 <BEV>  
A:Cross-references: UNIPROT:O49546; EMBL:AL021684  
A:Experimental source: cultivar Columbia; BAC clone F6H11  
C:Genetics:  
A:Map position: 5  
A:Introns: 54/2; 99/2; 134/2; 198/3; 235/3; 265/2; 309/2; 360/2; 395/2; 420/3; 445/3; 48  
A:Note: F6H11.180

Query Match 12.1%; Score 354.5; DB 2; Length 746;  
Best Local Similarity 24.5%; Pred. No. 1.8e-16;  
Matches 119; Conservative 72; Mismatches 185; Indels 109; Gaps 15;

Qy 107 ILSLGNVNLDRYTVTA-----GVLV-----LDIQHFGVDRGAGLIQSIFIC 149  
Db 269 ILCINILNIVDRGVIAASNGVSSKVCDAKGVCSAGTIGQGFNLTFEDGLLSAPMV 328  
Qy 150 SFMVAAPFVGLGDRNRK-----VILSCGIF-----FWSAVTFSSSFIPOQY 192  
Db 329 GLLVASPIFAGLSKRFNYQQHFVFLFFGVFNPFKILGVGLTVWTIAVIGCGF--SYN 386  
Qy 193 FWLLVLSRGLVGIGSEASTIAPTIIIGDLFTKNTLMSVFPFALPLSGSLGYITGSSV 252  
Db 387 FWMIAVFRMFVGEASEAFISLAAPYIDDSAPARKNFWLGLFYNCIPAGVALGVVFGYI 446  
Qy 253 KQAGDWHALRVSPVLGMITGTILILVLPATK-RGHADQ-----LGD 294  
Db 447 GNHLG-WRWAFYIEAIAMAVFVILSCFCIKPPQQLKGFADKSKKPSSTIETVAPTDAEAS 505

Qy 295 QLKARTS-----WLRDMKALIRNRSYVFPSSLSATSAVSPATGALGMWIP-----LYLH 341  
Db 506 QIKTKTPKSKNVLVFLGKDLKALPSEKVFIVNVGLGYITYNFVIGAYSYWGPKAGFIY-- 563  
Qy 342 RAQVQKTAETCNSPPCGAKDSLIFGAICTFTGFLGVVGTGAGATWCKLTQADPLVCA 401  
Db 564 ----KMKNAD-----MIFGGLTIICGIIIGLGGSYVLDNRINATLSNTFKLDA 607  
Qy 402 VGMGSAIFCLIFVAAKSSIVGAYICIF-VGETLLFSNMAITADILMYVVIPTRRATAV 460  
Db 608 STLGA-----FCTAFMKMNAFIALFAVEGELIFAPQAPVNFVCLHCVNLRPLSM 663  
Qy 461 ALOSFTSHLGDAGSPYLIGFISDLIRQSTKDSPLWFLSLGYALMLCPFVVVLGGMFFL 520  
Db 664 ASSTVLHILGDPSSPLYGKMDHLKWRKST-----LILTSILFL 705  
Qy 521 ATALP 525  
Db 706 AAIW 710

RESULT 6  
G87290  
major facilitator family transporter CC0336 [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C:Accession: G87290  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolc  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.;  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: G87290  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-438 <STO>  
A:Cross-references: UNIPROT:Q9AB95; GB:AE005673; NID:gl3421483; PIDN:AAK22323.1; GSPDB  
C:Genetics:  
A:Gene: CC0336

Query Match 9.8%; Score 285; DB 2; Length 438;  
Best Local Similarity 25.3%; Pred. No. 5.2e-12;  
Matches 112; Conservative 73; Mismatches 220; Indels 38; Gaps 14;

Qy 93 KPASLRG-----RGAAAAILSLGNVNLDRYTVAGVLL-DIQHFGVDRGAGLIQSIF 147  
Db 12 KAETAGRGGRYRVVLAAMLILAYTFNFLDR-QILGILAGSIKAEHLHLDLTDQLGLMGVAF 70  
Qy 148 ICSFMVAAPFVGLGDRFNKVLISCGIFPWSAVTFSSSFIPOQYFVLLVLSRGLVGIGE 207  
Db 71 AALYTLGVPLAMLADRVKSTWIMTVALTVWSGFTVVCGLAGG--FWSLFLARMGVGVE 128  
Qy 208 ASYSTIAPTIIIGDLFTKNTLMSVFPFALPLSGSLGYITGSSVKAAGDWHALRVSP 267  
Db 129 AGGVAPAYSILADYFPEQRARALAVVSGFPLGTALGVLFGLI-AAYVDWRFAFVAVG 187  
Qy 268 VLGMITGTLILVLPATKRGHADQL-GDQKAR-----TSWLRDMKALIRNRSYVFPSSLSAT 322  
Db 188 LAGVAFAPIFKVVVKOPVRGLDRAFGVAPAPPPKAPAFGQVLATVLPKPSFWLLSFGA 247  
Qy 323 SAVSPATGALGMWIPLYLHRAQVQKTAETCNSPPCGAKDSLIFGAICTFTGFLGVVTVGA 382  
Db 248 ACSICGCVAFVWLPFTFFQSFGLSLTDR-----ALYTSALUSLFGGVAVIGWG 296  
Qy 383 GATWCKLTQADPLVCAVGMGSAIFCLIFVAAKSSIVGAYICIFV--GETLLFSN 440  
Db 297 VLADRGAKNAAYALAPICFL--VALPCFLLAMNVQSMVAFLFLIPTGLNLAMLG 354  
Qy 441 AITADILMYVVIPTRRATAVALQSFTHSLGDAGSPYLIGFISDLIRQSTKDSPLWFLS 500  
Db 355 VWAA--VQHLAPPSMRTTTSALFLLINLLGLAVGLWFFGVVSDLL-----TPRYGAE 406

QY 501 LCVLWML-CPFVVVLGGMFFLAT 522  
Db 407 MRYAIYYGLSFYVVAVLLILAS 429

RESULT 7

F83064  
probable MFS transporter PA4654 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C:Accession: F83064  
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
A:Reference number: A82950, MUID:20437337, PMID:10984043  
A:Accession: F83064  
A:Molecule type: DNA  
A:Status: preliminary  
A:Residues: 1-457 <STO>  
A:Cross-references: UNIPROT:Q9HVD8; GB:AE004879; GB:AE004091; NID:g9950901; PIDN:AAG0804  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA4654

Query Match 9.5%; Score 278; DB 2; Length 457;  
Best Local Similarity 23.7%; Pred. No. 1.6e-11;  
Matches 108; Conservative 77; Mismatches 185; Indels 86; Gaps 16;  
QY 106 AILSLGNVLYDRYTVAGVLLDIQOHFGVKDGRGAGLQSVFICSPWVAAP1-FGVLDGR 164  
Db 19 AILWAVLYSFIDRQITLNLVGFIRKDLAISDTMSLMGLSFALPYTCGIPLGRMAN 78  
QY 165 FNRKVLSCGIFWSAVTFSSFIPOQYFWLVLGRGLVGEASYSTIAPTIIGDLFTK 224  
Db 79 RSRGILFLGVLSAMTAACGL--ARSYQWFLTRFVGVGVEALSPAYSIIADSFR 136  
QY 225 NTRTLMVYFPAIFLGSLGYITGSSV---KQAGDMH-----WAL--RVSPVLG 270  
Db 137 ERRATAISVYSGIYGLGSLAFLLGLVVKFASQAGDVHLPLFGEVRPQLIFLILGAAG 196  
QY 271 MITGTILILVPAIKRGHADQDQLKARTSWLRDMKALIRNSYVFSSLAISAVSFATG 330  
Db 197 VLFCLLLLAIREPARGVGAGVAVPLGVEGYALRANRKTVLCHNFGFACL-----SFAGY 251  
QY 331 ALGMMIPLYLHRAQVQKTAETCNSPPCGAKDSLIFGAICTCTGFLGVVTGAG-ATRWCR 389  
Db 252 GSGAWPTFFVRTHGWD-----AGHVGVVYGSIVAVFGCLGI VFGGLADYWA- 299  
QY 390 LKTQRADPLVCVGMGLSAIFICLIPVAAKSSIVGAYICIFVGETLLF-----SNW--AI 442  
Db 300 -KRGSDANN-RVGLL--AAWAVIP-----TLVYPLLDNANWAAAL 337  
QY 443 TADILMVVIP-----TTRATAVALQSTSHLLGDAGSPVLI GFISDLIRQ 488  
Db 338 MAPTVFELSPFGVAPAAIQEIMPNSMRGQASAIYLVVTLFGLGVGPTAVALTVDV-- 395  
QY 489 STKDSLWELFLSLGYALMLCPFVVVLGGMFFLATL 524  
Db 396 -----FADDMALRYSLLLVTLAANLVGAVVLIGL 425

RESULT 8

E87557  
major facilitator family transporter CC2486 [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C:Accession: E87557  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonin, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: E87557  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-519 <STO>  
A:Cross-references: UNIPROT:Q9A5G4; GB:AE005673; NID:g13424039; PIDN:AAK24457.1; GSPDB:1C  
C:Genetics:  
A:Gene: CC2486

Query Match 9.5%; Score 276; DB 2; Length 519;  
Best Local Similarity 24.3%; Pred. No. 2.5e-11;  
Matches 111; Conservative 71; Mismatches 188; Indels 86; Gaps 14;  
QY 82 ATAKPGGAQPKPASPGLRGRGAAAILSL---GNVLNLYDRYTVAGVLLDIQOHFGVKDR 138  
Db 24 AENHASGGDRPLYSN-----GYKATVLLGLLATYTFNFIDRTIISTIGOAIVDKLTD 78  
QY 139 GAGLLQSVFICSEFWVAAPIFGV---LGDENRNVILSCGIFWFAVTFSSSFIPOQYFW 194  
Db 79 QLGLLGGY---FALLYTLIGIPIARLAERFNRVTIISVSLVINSG--FTALCGAAANPA 133  
QY 195 LLVLSRGLVGEASYSTIAPTIIGDLFTKNTRTLMLSVFVPAIFLGSLGYITGSSVKQ 254  
Db 134 QLALFRFGVGEAGCSPSHSLISDYEPKRTALSIYSGIPLGTWFGAVAGWLQAQ 193  
QY 255 AAGDWHWALRVSPLVGMITGTLLILVPAIKRGHA-----DQLGDQLKARTSWLR 304  
Db 194 EP-SNRFAFVIGLPGVILLAVIKLVKVPKPRGSEIVERPLEAEDVVVEPAKPAFMSMAN 252  
QY 305 DMKALIRNSYVFSS-----LATSAVSFATGALGMWIPLYLHRAQVQKTAETCNSPP 357  
Db 253 EFKELWAVTKILFGKWPVLMVGLVGTIASFGAYSGAFVPSYFVRAFDLGL----- 303  
QY 358 CGAKDSLIFGAICTCTGFLGVVTGAGATRWCLRTQRADPLVCVGMGLSAIFICL-IPV 416  
Db 304 --AQVGLITGLIGFSAGVGLVGGFLSDWAGRSKAKVALTPAIGLI-----LCTPIVI 356  
QY 417 AAKSSIVGAYICIFVGETLLFSNWAITADILM-----YV-----VIPRRAT 458  
Db 357 LA-----YLOQDMQTALLILVPGIFHYVVLAPTFFGVVQNSVEPRRRAT 400  
QY 459 AVALQSTSHLLGDAGSPVLI GFISDLIRQSTKQSP 494  
Db 401 ATALLFFFLNIALGVGPFVFTGWLIDHLAQHFNNP 436

RESULT 9

F65250  
hypothetical 49.4 kD protein in ter-mdoB intergenic region - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: F65250; S56582  
R:Blattner, F.R.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.  
A:Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: F65250  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-453 <BLAT>  
A:Cross-references: UNIPROT:P39398; GB:AE000506; GB:U00096; NID:g2367377; PIDN:AAC77312  
A:Experimental source: strain K-12, substrain MG1655  
R:Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.  
Nucleic Acids Res. 23, 2105-2119, 1995  
A:Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92  
A:Reference number: S56344; MUID:95334362; PMID:7610040  
A:Accession: S56582  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 262-453 <BUR>

A;Cross-references: EMBL:U14003; NID:g1263172; PIDN:AAA97253.1; PID:g537198  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994  
C;Superfamily: hexuronate transporter

Query Match 8.6%; Score 252; DB 1; Length 453;  
Best Local Similarity 24.3%; Pred. No. 9.3e-10;  
Matches 115; Conservative 79; Mismatches 212; Indels 68; Gaps 19;

```
Qy 72 PGTGPTGCAATAGP--GAQPKPASLGRGRGAAAILSLGNVNLNLDRTYVAGVLLDI 129
Db 10 PRSFTSSSADIPVPDGLVQ-RSTRIKRIQTAMLLFFAAVINYLDRSSLSVANLTI 68

Qy 130 QQHFGVKDRGAGLLQSVFCFMSVVAAPIFGLGRDNKRVILSCGIFFWSAVTSSFP 189
Db 69 REELGLSATIGALLSVFSLAYGIAQLPCGPLLDRKGRPMGLGMFFWLSLFOAMSGMV- 127

Qy 190 QQYFWLLVLSRGLVIGEASYSTIAPTIIGDLFTKNTRTLMLSVFYFAIPLGSLGYITG 249
Db 128 -HSFTQFVLVRIGWGIGEAPMPCGVKVINDFWNIKERGRPMGFNAASTIGVAVSPPI 186

Qy 250 SSVKQAAGDWHWALRVSPVLG--MITGLTIL-----ILVPATKRGHADQLGDOLKARTS 301
Db 187 AAMMLVNG-WRGMPFITIGVLGIFLAIGWYMLYRNREHVELTAVEQAYLN--AGSVNARD 243

Qy 302 WL--RDMKALIRNSYVFSLSATSAVSFATGALGMWIPLYLHRAQVQVKTAEKTCNSPPCG 359
Db 244 PLGFAEWRSLFRNRTMGMMLGFGSGINYTAWLPLGVLQTAYNLD----- 291

Qy 360 AKDSLIFGATCTGTFGLGVVGTAGATRW-----CRKTKORADPLVCVAV-GMLGSAIFIC 412
Db 292 LKSTGLMAAIPFLFGAAGMLVNGVYTDWLVKGGMAPIKSRK----ICIIAGMFCSAATL 347

Qy 413 LIFVAAKSS-----IVG-AYICIFVGETLLFSNWAITADILMYVVIPTRRATAV-ALQSET 466
Db 348 IVPQATTSMTAVLLIGNALFCIHFGAT---SCWG-----LIHVAVASRMTASVGSIQNFA 399

Qy 467 SHLLGDAGSPYLIGFISDLIRQSTKDSPLWEFLSLGYALMLCPFVVVLGGMFFL 520
Db 400 SFICA-SFAPITGTGFIVDTH-----SFRLLAIICGCVTAAGALAYI 440
```

RESULT 10  
D91293  
probable transport protein Ecs5316 [imported] - Escherichia coli (strain O157:H7, substra  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C;Accession: D91293  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuwara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: D91293  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-453 <RAY>  
A;Cross-references: UNIPROT:Q8XEG8; GB:BA000007; PIDN:BA38739.1; PID:g13364794; GSPDB:G  
A;Experimental source: strain O157:H7, substrain RIMD 0509952  
C;Genetics:  
C;Superfamily: hexuronate transporter

Query Match 8.6%; Score 252; DB 2; Length 453;  
Best Local Similarity 24.3%; Pred. No. 9.3e-10;  
Matches 115; Conservative 79; Mismatches 212; Indels 68; Gaps 19;

```
Qy 72 PGTGPTGCAATAGP--GAQPKPASLGRGRGAAAILSLGNVNLNLDRTYVAGVLLDI 129
Db 10 PRSFTSSSADIPVPDGLVQ-RSTRIKRIQTAMLLFFAAVINYLDRSSLSVANLTI 68

Qy 130 QQHFGVKDRGAGLLQSVFCFMSVVAAPIFGLGRDNKRVILSCGIFFWSAVTSSFP 189
Db 69 REELGLSATIGALLSVFSLAYGIAQLPCGPLLDRKGRPMGLGMFFWLSLFOAMSGMV- 127
```

```
Qy 190 QQYFWLLVLSRGLVIGEASYSTIAPTIIGDLFTKNTRTLMLSVFYFAIPLGSLGYITG 249
Db 128 -HSFTQFVLVRIGWGIGEAPMPCGVKVINDFWNIKERGRPMGFNAASTIGVAVSPPI 186

Qy 250 SSVKQAAGDWHWALRVSPVLG--MITGLTIL-----ILVPATKRGHADQLGDOLKARTS 301
Db 187 AAMMLVNG-WRGMPFITIGVLGIFLAIGWYMLYRNREHVELTAVEQAYLN--AGSVNARD 243

Qy 302 WL--RDMKALIRNSYVFSLSATSAVSFATGALGMWIPLYLHRAQVQVKTAEKTCNSPPCG 359
Db 244 PLGFAEWRSLFRNRTMGMMLGFGSGINYTAWLPLGVLQTAYNLD----- 291

Qy 360 AKDSLIFGATCTGTFGLGVVGTAGATRW-----CRKTKORADPLVCVAV-GMLGSAIFIC 412
Db 292 LKSTGLMAAIPFLFGAAGMLVNGVYTDWLVKGGMAPIKSRK----ICIIAGMFCSAATL 347

Qy 413 LIFVAAKSS-----IVG-AYICIFVGETLLFSNWAITADILMYVVIPTRRATAV-ALQSET 466
Db 348 IVPQATTSMTAVLLIGNALFCIHFGAT---SCWG-----LIHVAVASRMTASVGSIQNFA 399

Qy 467 SHLLGDAGSPYLIGFISDLIRQSTKDSPLWEFLSLGYALMLCPFVVVLGGMFFL 520
Db 400 SFICA-SFAPITGTGFIVDTH-----SFRLLAIICGCVTAAGALAYI 440
```

RESULT 11

G86134  
hypothetical protein yjiz [imported] - Escherichia coli (strain O157:H7, substrain EDL9  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: G86134  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimmalanta, E.; Potamousis, K.; Apodaca  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: G86134  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-453 <STO>  
A;Cross-references: UNIPROT:Q8XEG8; GB:AB005174; NID:g12519379; PIDN:AAG59539.1; GSPDB:  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
C;Gene: Yjiz  
C;Superfamily: hexuronate transporter

Query Match 8.6%; Score 252; DB 2; Length 453;  
Best Local Similarity 24.3%; Pred. No. 9.3e-10;  
Matches 115; Conservative 79; Mismatches 212; Indels 68; Gaps 19;

```
Qy 72 PGTGPTGCAATAGP--GAQPKPASLGRGRGAAAILSLGNVNLNLDRTYVAGVLLDI 129
Db 10 PRSFTSSSADIPVPDGLVQ-RSTRIKRIQTAMLLFFAAVINYLDRSSLSVANLTI 68

Qy 130 QQHFGVKDRGAGLLQSVFCFMSVVAAPIFGLGRDNKRVILSCGIFFWSAVTSSFP 189
Db 69 REELGLSATIGALLSVFSLAYGIAQLPCGPLLDRKGRPMGLGMFFWLSLFOAMSGMV- 127

Qy 190 QQYFWLLVLSRGLVIGEASYSTIAPTIIGDLFTKNTRTLMLSVFYFAIPLGSLGYITG 249
Db 128 -HSFTQFVLVRIGWGIGEAPMPCGVKVINDFWNIKERGRPMGFNAASTIGVAVSPPI 186

Qy 250 SSVKQAAGDWHWALRVSPVLG--MITGLTIL-----ILVPATKRGHADQLGDOLKARTS 301
Db 187 AAMMLVNG-WRGMPFITIGVLGIFLAIGWYMLYRNREHVELTAVEQAYLN--AGSVNARD 243

Qy 302 WL--RDMKALIRNSYVFSLSATSAVSFATGALGMWIPLYLHRAQVQVKTAEKTCNSPPCG 359
Db 244 PLGFAEWRSLFRNRTMGMMLGFGSGINYTAWLPLGVLQTAYNLD----- 291

Qy 360 AKDSLIFGATCTGTFGLGVVGTAGATRW-----CRKTKORADPLVCVAV-GMLGSAIFIC 412
```

Db 292 LKSTGLMAIPFLFGAAGMLVNGVYTDWLKVGGMAPIKSRK----ICIIAGMFCSAAFTL 347

Qy 413 LIFVAAKSS-----IVG-AYICIFVGETLLFSNWAITADILMYVVIPTERRATAV-ALQSFT 466

Db 348 IVPQATTSMAVLLIGMALFCIIFAGT-----SCWG-----LIHVAVASRMTASVGSIQNFA 399

Qy 467 SHLLGDAGSPYLIGFISDLIRQSTKDSPLMEFLSLGVALMLCPVVVVLGGMFFL 520

Db 400 SFICA-SFAPITIGFIVDTH-----SFRLLAIICGCVTAAGALAYI 440

RESULT 12

Db7557

major facilitator family transporter CC2485 [imported] - Caulobacter crescentus

C/Species: Caulobacter crescentus

C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004

C/Accession: Db7557

R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonitskii, N. J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A/Title: Complete Genome Sequence of *Caulobacter crescentus*.

A/Reference number: A87249; MUID:21173698; PMID:11259647

A/Accession: Db7557

A/Status: preliminary

A:Molecule type: DNA

A:Residues: 1-444 <STO>

A/Cross-references: UNIPROT:Q9A5G5; GB:AE005673; NID:g13424038; PIDN:AAK24456.1; GSPDB:G

C:Genetics:

A:Gene: CC2485



This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 11:01:18 ; Search time 111 Seconds  
(without alignments)  
2933.892 Million cell updates/sec

Title: US-10-085-198-48  
Perfect score: 2920  
Sequence: 1 MMCLCASAAGGAEEED.....VHQRGPGCTALAHVRVGAS 566

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02.\*

1: uniprot\_sprot.\*

2: uniprot\_treml.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2761	94.6	619	2 Q81VW8	Q81VW8 homo sapien
2	2652	90.8	590	2 Q91VM4	Q91VM4 mus musculus
3	1298	44.5	528	2 Q9H2V7	Q9H2V7 homo sapien
4	1297	44.4	524	2 Q7L541	Q7L541 homo sapien
5	1297	44.4	528	2 Q6F182	Q6F182 homo sapien
6	1297	44.4	528	2 AAH65235	AAH65235 homo sapi
7	1289	44.1	528	2 Q8R0G7	Q8R0G7 mus musculus
8	1278	43.8	528	2 Q9EOK0	Q9EOK0 mus musculus
9	1266.5	43.4	506	2 Q8UVB7	Q8UVB7 brachydanio
10	1264.5	43.3	506	2 Q7ZU13	Q7ZU13 brachydanio
11	1258.5	43.1	486	2 Q99LN7	Q99LN7 mus musculus
12	1167	40.0	476	2 Q86VU7	Q86VU7 homo sapien
13	1139	39.0	514	2 Q9D232	Q9D232 mus musculus
14	1115	38.2	420	2 Q71RB5	Q71RB5 homo sapien
15	1115	38.2	420	2 AAQ15259	AAQ15259 homo sapi
16	1088	37.3	454	2 Q8TCS5	Q8TCS5 homo sapien
17	1080	37.0	512	2 Q6ZMD2	Q6ZMD2 homo sapien
18	1080	37.0	512	2 BAD18797	BAD18797 homo sapi
19	1049	35.9	492	2 Q8BYA9	Q8BYA9 mus musculus
20	965	33.0	643	2 Q7Q5N4	Q7Q5N4 anopheles g
21	964.5	33.0	605	2 Q9GQ00	Q9GQ00 drosophila
22	959	32.8	630	2 Q9GQ02	Q9GQ02 drosophila
23	930.5	31.9	605	2 Q9GQ01	Q9GQ01 drosophila
24	925	31.7	630	2 Q9GQ03	Q9GQ03 drosophila
25	916.5	31.4	477	2 Q960X6	Q960X6 drosophila
26	874.5	29.9	518	2 Q9NEW5	Q9NEW5 caenorhabdi
27	855	29.3	531	2 Q01927	Q01927 caenorhabdi
28	816	27.9	385	2 Q81Z31	Q81Z31 homo sapien
29	815	27.9	488	2 Q18545	Q18545 caenorhabdi
30	710	24.3	401	2 Q19235	Q19235 caenorhabdi
31	644	22.1	229	2 Q8N953	Q8N953 homo sapien

32 633.5 21.7 402 2 Q8MLD0 Q8mid0 drosophila  
33 633.5 21.7 402 2 Q9GQP9 Q9GQP9 drosophila  
34 442 15.1 118 2 Q8R119 Q8R119 mus musculus  
35 392.5 13.4 507 2 Q9ZQ41 Q9ZQ41 arabidopsis  
36 380 13.0 492 2 Q6NMN6 Q6NMN6 arabidopsis  
37 380 13.0 492 2 AAS47627 AAS47627 arabidops  
38 380 13.0 492 2 AAT41792 AAT41792 arabidops  
39 355 12.2 484 2 Q9FLG8 Q9FLG8 arabidopsis  
40 354.5 12.1 746 2 Q49546 Q49546 arabidopsis  
41 327 11.2 471 2 Q6J2D6 Q6J2D6 pseudomonas  
42 327 11.2 471 2 Q88BM0 Q88BM0 pseudomonas  
43 327 11.2 471 2 AAT35185 AAT35185 pseudomon  
44 324 11.1 479 2 Q87XS1 Q87x81 pseudomonas  
45 319.5 10.9 431 2 Q89KI2 Q89ki2 bradyrhizob

ALIGNMENTS

RESULT 1

Q81VW8 PRELIMINARY; PRT; 619 AA.  
AC Q81VW8  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Similar to spinster-like protein (Fragment).  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
DR EMBL; BC041772; AAH41772.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR007114; MFS.  
DR Pfam; PF00803; Sugar tr; 1.  
DR PROSITE; PS50850; MFS; 1.  
KW Transmembrane.  
FT NON\_TER  
SQ SEQUENCE 619 AA; 64767 MW; 77C223EB281C4286 CRC64;

Query Match 94.6%; Score 2761; DB 2; Length 619;  
Best Local Similarity 99.3%; Pred. No. 7, 1e-161;  
Matches 535; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MMCLCASAAGGAEEEDADERRRRRRRGAQAGAGCGCGAGAGGAGVSAAGDEVQTL 60  
Db 71 MMCLCASAAGGAEEEDADERRRRRRRGAQAGAGCGCGAGAGGAGVSAAGDEVQTL 130  
QY 61 SSVVRATPGPTGTGCAATAGCAQPKASLGRGCGAAAIISLGNVLNLYDRY 120  
Db 131 SSVVRATPGPTGTGCAATAGCAQPKASLGRGCGAAAIISLGNVLNLYDRY 190  
QY 121 TVAGVLLDIQQHFGVKDRGAGLLQSVFICSFVVAAPIFGYLGDNRNKRKVLSCGIFFWSA 180  
Db 191 TVAGVLLDIQQHFGVKDRGAGLLQSVFICSFVVAAPIFGYLGDNRNKRKVLSCGIFFWSA 250  
QY 181 VTFSSSFIQQYFWLLVLSRGLVIGEASYSYTIPTIIGDLFTKNTRTLMLSVFFFAIPL 240  
Db 251 VTFSSSFIQQYFWLLVLSRGLVIGEASYSYTIPTIIGDLFTKNTRTLMLSVFFFAIPL 310  
QY 241 GSGLYITGSSVKQAAGDWHALRVSPVLGMITGTLLILVLPATKRGHADQIGDQKART 300  
Db 311 GSGLYITGSSVKQAAGDWHALRVSPVLGMITGTLLILVLPATKRGHADQIGDQKART 370

```
Qy 301 SWLRDMKALIRNRSYVFSSLSATSAVSFATGALGMWIPLYLHRAQVQKTAETCNPPCGA 360
Dy 371 SWLRDMKALIRNRSYVFSSLSATSAVSFATGALGMWIPLYLHRAQVQKTAETCNPPCGA 430
Qy 361 KDSLIIFGAICTFTGFLGVVTCAGATRCRLKTQADPLVCAVGMGLGSAIFCLIFVAAS 420
Dy 431 KDSLIIFGAICTFTGFLGVVTCAGATRCRLKTQADPLVCAVGMGLGSAIFCLIFVAAS 490
Qy 421 SIVGAYICIFVGETLLFSNWAITADILMVYVIPTRRATAVALQSFTHLLGDAGSPYLIG 480
Dy 491 SIVGAYICIFVGETLLFSNWAITADILMVYVIPTRRATAVALQSFTHLLGDAGSPYLIG 550
Qy 481 FISDLIRQSTKDSPLWEFLSLGYALMLCPFPVVVLGGMFFLATALFFVSDRARAEOHIGE 539
Dy 551 FISDLIRQSTKDSPLWEFLSLGYALMLCPFPVVVLGGMFFLATALFFVSDRARAEOHIGE 609

RESULT 2
Q91VM4 ID Q91VM4 PRELIMINARY; PRT; 590 AA.
AC Q91VM4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BC011467 protein (Fragment).
GN Name=BC011467;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=129; TISSUE=Mammary tumor. Brcal-1/fl;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci F., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Woxley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=129; TISSUE=Mammary tumor. Brcal-1/fl;
RX Strausberg R.;
RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; BC011467; AAH11467.1; -
DR MGD; MGI:2384936; BC011467.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PS50850; MFS; 1.
DR Transmembrane.1
FT NON_TER
SQ SEQUENCE 590 AA; 61956 MW; 029412B949E1E0CF CRC64;
```

```
Query Match 90.8%; Score 2652; DB 2; Length 590;
Best Local Similarity 95.4%; Pred. No. 3.2e-154;
Matches 514; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

Qy 1 MMCLCASAAGCAEEDAEERRRRRGNORAGSGCCGARGAGAGYSAAGDEVOTL 60
Dy 42 MMCLCASAAGCAEEDAEERRRRRGNORAGSGCCGARGAGAGYSAAGDEVOTL 101
Qy 61 SGSVRRAPTPGPGCAATAGKGPAAQPKPASLGRGRGAAAILSLGNVLYLDYR 120
Dy 102 SGSVRRAPTPGPGCAATAGKGPAAQPKPASLGRGRGAAAILSLGNVLYLDYR 161
Qy 121 TVAGVLLDIQHFQVGRDAGLLQSVFICSFVMAAPIFGYLGRDRNRKVLSCGIPFWSA 180
Dy 162 TVAGVLLDIQHFQVGRDAGLLQSVFICSFVMAAPIFGYLGRDRNRKVLSCGIPFWSA 221
Qy 181 VTFSSSFIPOQYFWLLVLSRGLVIGIEASYSTIAPTIIGDLFTKNTTLMLSVYFAIPL 240
Dy 222 VTFSSSFIPOQYFWLLVLSRGLVIGIEASYSTIAPTIIGDLFTKNTTLMLSVYFAIPL 281
Qy 241 GSGLGVIITGSSVKQAAGDHWALRVSPVLGMITGTLILVLPATKRGHADQLGQDKART 300
Dy 282 GSGLGVIITGSSVKQAAGDHWALRVSPVLGMITGTLILVLPATKRGHADQLGQDKART 341
Qy 301 SWLRDMKALIRNRSYVFSSLSATSAVSFATGALGMWIPLYLHRAQVQKTAETCNPPCGA 360
Dy 342 SWLRDMKALIRNRSYVFSSLSATSAVSFATGALGMWIPLYLHRAQVQKTAETCNPPCGA 401
Qy 361 KDSLIIFGAICTFTGFLGVVTCAGATRCRLKTQADPLVCAVGMGLGSAIFCLIFVAAS 420
Dy 402 KDSLIIFGAICTFTGFLGVVTCAGATRCRLKTQADPLVCAVGMGLGSAIFCLIFVAAS 461
Qy 421 SIVGAYICIFVGETLLFSNWAITADILMVYVIPTRRATAVALQSFTHLLGDAGSPYLIG 480
Dy 462 SIVGAYICIFVGETLLFSNWAITADILMVYVIPTRRATAVALQSFTHLLGDAGSPYLIG 521
Qy 481 FISDLIRQSTKDSPLWEFLSLGYALMLCPFPVVVLGGMFFLATALFFVSDRARAEOHIGE 539
Dy 522 FISDLIRQSTKDSPLWEFLSLGYALMLCPFPVVVLGGMFFLATALFFVSDRARAEOHIGE 580

RESULT 3
Q9H2V7 ID Q9H2V7 PRELIMINARY; PRT; 528 AA.
AC Q9H2V7; Q9BRN5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Spinster (SPINL protein).
GN Name=SPINL;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=21238517; PubMed=11340170;
RA Nakano Y., Fujitani K., Kurihara J., Ragan J., Ueui-Aoki K.,
RA Shimoda L., Lukacovich T., Suzuki K., Sezaki M., Sano Y., Ueda R.,
RA Awano W., Kaneda M., Umeda M., Yamamoto D.;
RT "Mutations in the novel membrane protein spinster interfere with
RT programmed cell death and cause neural degeneration in Drosophila
RT melanogaster.";
RL Mol. Cell. Biol. 21:3775-3788 (2001).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci F., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Woxley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=129; TISSUE=Mammary tumor. Brcal-1/fl;
RX Strausberg R.;
RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; BC011467; AAH11467.1; -
DR MGD; MGI:2384936; BC011467.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PS50850; MFS; 1.
DR Transmembrane.1
FT NON_TER
SQ SEQUENCE 590 AA; 61956 MW; 029412B949E1E0CF CRC64;
```

Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;  
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[3]  
SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
[4]  
SEQUENCE FROM N.A.  
RC TISSUE=Pancreas;  
RA Strausberg R.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
EMBL: AF212371; AAG43830.1; -;  
DR EMBL; BC008325; AAH08325.1; -;  
DR EMBL; BC038961; AAH38961.1; -;  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005215; P:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub transporter.  
DR Pfam; PF00083; Sugar tr; 1.  
DR PROSITE; PS00850; MFS; 1.  
KW Transmembrane.  
SQ SEQUENCE 528 AA; 56629 MW; F1B9D2EB3F9F1B48 CRC64;  
Query Match 44.5%; Score 1298; DB 2; Length 528;  
Best Local Similarity 53.4%; Pred. No. 1.9e-71;  
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;  
QY 53 AGDEVQTLGSGVRRAPTGE-PGTGPTGCAATAGKGAQGPASLGR-----GRGA-A 104  
DB 2 AGSDTAPFLSQADDDPDGPGVGTGPGTGNPKSEEPVQEGQLRITGLSPGESALI 61  
QY 105 AAILSGVNLVLDRTYVAGVLLDIOQHFQVKGDRGAGLQSVFICSFVMAAPFVGLGDR 164  
DB 62 VAVLCVNLNLYMDRTFVAGVLPDIEQFFNIGDSSGLITQTVFISSYMWLAPVFGVGLDR 121  
QY 165 FNRKVLSCGIFPWSAVTSSFPQQYFWLLVSLRGVNGEASVSTIAPTIIIGDLPTK 224  
DB 122 YNRKYLKMGCGIAFWSLVTLTGSSFPGEHFVLLLTGLVGVGEASVSTIAPTIIADLFVA 181  
QY 225 NRTLMLVFPYFAIPGSLGTYTGSSVKQAGDHWALRVSPVLCMIGTGLILVLPAT 284  
DB 182 DQSRMLSLTFYFAIPVGSGLGVTAGSKVKDMAGDHWALRVTPGLGVAVLLFLVVRP 241  
QY 285 KRGHADQDQLKAR-TSWLRDMKALIRNRSVYFSSLATSAVSFATGALGMWIPYLHRA 343  
DB 242 PRGAVRHSDLPNLTNSWADLRALARNPSFVLSLGTFAVAVFTGSLALWAPAFLLRS 301  
QY 344 QVYQKTAETC-NSPPCGAKDSLIFGATCTFTGFLGVVTGAGATRCRLKATQADPLVCV 402  
DB 302 RVVLGETPPLCPGDCSSSDSLIFGLITCLTGLVGLGVGEISVRLRHSNPRADPLVCAT 361  
QY 403 GMLGSAIFCLIFVAAKSSIVGAYICIFVGETLLFSNWAITADILMYVYIPTRRTATAVAL 462  
DB 362 GLGGAPPLFSLACARGIVATYPIFIFGETLLSNWVAIVADILLYVYIPTRRTAEAF 421  
QY 463 QSTSHLLDAGSPYLIGFISDLIRQSTKDSPLWEFLSLGYALMLCPFFVVLGGMFFLAT 522

Db 422 QIVLSHLGDSAGSPYLIGLISDLRLRNWPPSPFLSEFRALQFSLMLCAFVGALGGAFLGT 481  
QY 523 ALFFVSDRARABOHLGE--RRAGVR---VHQRG 551  
DB 482 AFIIEADRRRAHQHVGQLLHEAGSTDDRIIVPQRG 516  
RESULT 4  
Q7L541  
ID Q7L541 PRELIMINARY; PRT; 524 AA.  
AC Q7L541;  
DT 05-JUL-2004 (TREMBLrel. 27, Created)  
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
DE SPINL protein (Fragment).  
GN Name=SPINL;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;  
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
DR EMBL; BC006156; AAH06156.1; -;  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub transporter.  
DR Pfam; PF00083; Sugar tr; 1.  
DR PROSITE; PS00850; MFS; 1.  
KW Transmembrane.  
FT NON\_TER  
SQ SEQUENCE 524 AA; 56283 MW; CSEBA9DB206256E4 CRC64;  
Query Match 44.4%; Score 1297; DB 2; Length 524;  
Best Local Similarity 54.8%; Pred. No. 2.2e-71;  
Matches 273; Conservative 61; Mismatches 148; Indels 16; Gaps 7;  
QY 70 GP-PGTGPTGCAATAGKGAQGPASLGR-----GRGA-AAAILSGVNLVLDRTY 121  
DB 15 GPVPTGTLPGSTGNPKSEEPVQEGQLRITGLSPGESALIIVAVLCVNLNLYMDRPT 74  
QY 122 VAGVLLDIOQHFQVKGDRGAGLQSVFICSFVMAAPFVGLGDRNKRKVLSCGIFPWSAV 181  
DB 75 VAGVLPDIEQFFNIGDSSGLITQTVFISSYMWLAPVFGVGLGDRYNRKVLKMGCGIAFWSLV 134  
QY 182 TFSFSSFPQQYFWLLVSLRGVNGEASVSTIAPTIIIGDLPTKNTLMLSVFPYFAIPLG 241

```
Db 135 TLGSSFIPEGHFWLILLTRGLVGVGEASYSYTIPTLIADLFVADQSRMLSIIFYFAIPVG 194
Qy 242 SCLGYITGSSVKQAGDWHALRVPVGLMITGTLILLVLPATRGHADQLGDQKAR-T 300
Db 195 SCLGYIAGSKVDWAGDWHALRVPVGLVAVALLFLVREPRGAVRHSDDLPLNPT 254
Qy 301 SWLRDMKALIRNRSYVFSLSLTSVAVSFATGALGMWIPLYLHRAOVVQKTAETC-NSPPCG 359
Db 255 SSWADLRALARNPSPVLSLGLTAVAFVTGSLAWAPAFLLRSRVLGETPPCLPGDSCS 314
Qy 360 AKDSLIFGATCTCTGFLGVVGTGAGATRCWRUKTORADPLVCAVGMGLSAGIIFCLIFVAAK 419
Db 315 SSDLIFGLTCLTGLVGLGVEISRRLRSHNPRADPLVCAVGMGLSAGIIFCLIFLACAR 374
Qy 420 SSIIVGAYICIFVGETLPSNWAITAILMVYVPIPTRRATVALQSFTHLGDAGSPVLI 479
Db 375 GSVIATYIFIFIGETLLSMNNAIVADILLYVPIPTRRSTABAFQIVLSHLGDAGSPVLI 434
Qy 480 GFISDLIRQSKDPSLWFLSGLVYALMLCPFFVVLGGMFFLATALFFVSDRRARAEQHIGE 539
Db 435 GLISDLRRNWPSPFLSEFRALQSLMLCAFVGALGGAFLGTAIFIEADRRRAQLHVQG 494
Qy 540 --RRAGVR-----VVHQRG 551
Db 495 LLHEAGSTDDRIWVPQRG 512

RESULT 5
Q6P182
ID Q6P182 PRELIMINARY; PRT; 528 AA.
AC Q6P182;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE SPINL protein.
GN Name=SPINL;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Prange C.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (by similarity).
DR EMBL; BC065235; LOC007114; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
```

```
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PS00850; MFS; 1.
KW Transmembrane.
SQ SEQUENCE 528 AA; 56655 MW; BC03B6BF2BCB03A6 CRC64;

Query Match 44.4%; Score 1297; DB 2; Length 528;
Best Local Similarity 53.4%; Pred. No. 2.2e-71;
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;

Qy 53 AGDEVQTLSSVRRATGP--PGTPTGPGCAATAGKGAQOQPKPASLGR-----GRGA-A 104
Db 2 AGSDTAPFLSQADDDPGVPVGTGPGSTGNPKSEEPVDPQEGQLRITGLSPGRSALI 61
Qy 105 AAILSGNLVNLDRYTVAGVLLDIOOHFGVKDAGGLQSVFICSMVAAPFIFYGLGDR 164
Db 62 VAVLCYNLNLNMDRTFVAGVLDIQQFNFIGNSSGLIQTVFISSVWVLAAPFIFYGLGDR 121
Qy 165 FNRKVLSCGIFPWSAVTFSSSFIPOQYFWLLVLSRLGVIGIGEASYSTIAPTIGDLFTK 224
Db 122 YNRKYLWCGGIAPWSLVTLGSSFIPEGHFWLLLTGLVGVGEASYSTIAPTILADLFVA 181
Qy 225 NTRTLMLSVFPFAIPGLSGGLGYITGSSVKQAGDWHALRVSPVLGMITGTLILLVPAT 284
Db 182 DQSRMLSIIFYFAIPVGSGLGYTAGSKVKDMAGDWHALRVTPCLGVVPLLFLVVREP 241
Qy 285 KRGHADQLGDQKAR--TSLRDMKALIRNRSYVFSLSLTSVAVSFATGALGMWIPLYLHRA 343
Db 242 PRGAVRHSDDLPLNPTSTWADRLARNPSPFLSSGLGTAVAFVTGSLAWAPAFLLRS 301
Qy 344 QVVQKTAETC-NSPPCGAKDSLIFGATICTPTGFLVVVGTGAGATRCWLKTKQRADPLVCAV 402
Db 302 RVVLGETPPCLPGDSCSSSDSLIFGLTCLTGLVGLGVGEISRRLRSHNPRADPLVCAV 361
Qy 403 GMLGSAIFCLIFVAAKSSIVGAYICIFVGETLPSNWAITAILMVYVPIPTRRATVAL 462
Db 362 GLGSAFFFLSLACARGSIATYIFIFIGETLLSMNNAIVADILLYVPIPTRRSTAEAF 421
Qy 463 QSFTHLGDAGSPYILGFTSDLIROSKDPSLWFLSGLVYALMLCPFFVVLGGMFFLAT 522
Db 422 QIVLSHLGDAGSPYILGFTSDLIROSKDPSLWFLSGLVYALMLCPFFVVLGGMFFLAT 481
Qy 523 ALFFVSDRARAEOHGLGE--RRAGVR-----VVHQRG 551
Db 482 AIFIEADRRRAQLHVQGLLHEAGSTDDRIWVPQRG 516

RESULT 6
AAH65235
ID AAH65235 PRELIMINARY; PRT; 528 AA.
AC AAH65235;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE SPINL protein.
GN SPINL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (by similarity).
DR EMBL; BC065235; LOC007114; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
```

```
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Blakesley A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley A.C., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RA Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC065235; AAH65235.1; -.
SQ SEQUENCE 528 AA; 56655 MW; BC03B6BF2BCB03A6 CRC64;

Query Match 44.4%; Score 1297; DB 2; Length 528;
Best Local Similarity 53.4%; Pred. No. 2.2e-71;
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;

QY 53 AGDEVQTLSSVRRAPTGP-PGTPGTCGAATAKPGCAQGPASLGR-----GRGA-A 104
DB 2 AGSDTAPFLSQADDDPDGPGFPGTGTGNGPKSEEPVQEGQRTIGLSRGSALI 61

QY 105 AAILSLGNVLYDRYTVAGVLLDIOQHFGVKDRGAGLLQSVFCSFMVAAPFVGLGR 164
DB 62 VAVLCYINLNYWDRFTVAGVLDIEQFFNIGDSSGLQTFVFSYVWVAPVFGVGLGR 121

QY 165 FNRKVLSCGIFPWSAVTFSSFPQOYFWLLVSLRGLVIGEASYSTIAPTIGDLFTK 224
DB 122 YNRKYLKMGCGIAPFWSLVTLGSSFPICGEHFHLLLTGLVGVGEASYSTIAPTILDLFA 181

QY 225 NTRTLMLSVYFAIPGLSGVLYTGSVVKQAGDHWALRVSPVLGMITGTLILVLPAT 284
DB 182 DQSRMLSIIFYFAIPVGSGLGIYAGSKVMDAGDHWALRVTPGLGVVPLLVVREP 241

QY 285 KRGHADQLGPKAR-TSWLRDMKALIRNSVYVFFSLATSAYSFATGALGMWIPVYLRHA 343
DB 242 PRGAVEHSDLPPLNFTSNWADLALARNPFSVLSSLGFTAVAFVFGSLALWAPFLRS 301

QY 344 QVQKTAETC-NSPPCGAKDSLIFGAICTFTGFLGVVTVGAGATRCWLKTRQADPLVCV 402
DB 302 RVVLGETPPCLPGDCSSSDSLIFGLITCLTGLVGLGVGEISRLRHSNPRADPLVCAT 361

QY 403 GMLGSAIFCLIFVAAKSIVGAYICIFVGETLLFSNWAITADILMYVVIPTTRATAVAL 462
DB 362 GLUGSAPFLFLSLACARGSVATYFIFIGETLLSNMVAIVADILLYVVIPTTRATAEAF 421

QY 463 QGFTSHLLGDAGSPYLLIGFISDLIROSTKDSPLWELSLGYALMLCPFVVVLGGMFFLAT 522
DB 422 QIVLSHLLGDAGSPYLLIGLISDLRLRNWPPSFLSEFRLAQFSLMLCAFVGLGGAFLGT 481

QY 523 ALFFVSDRARAEOHLGE--RRAGVR---VVHQRG 551
DB 482 AIFIEADRRRAQLHVQGLLHEAGSTDDRIVVVQRG 516

RESULT 7
Q8R0G7 PRELIMINARY; PRT; 528 AA.
AC Q8R0G7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Spinder.
GN Name=2210013K02Rik;
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; BC026854; AAH26854.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000566; Lipocin_cytFABP.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; Sugar tr. 1.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
DR PROSITE; PS00850; MFS; 1.
DR Transmembrane.
SQ SEQUENCE 528 AA; 56708 MW; DDAB448D7D2B9E17 CRC64;

Query Match 44.1%; Score 1289; DB 2; Length 528;
Best Local Similarity 52.5%; Pred. No. 6.8e-71;
Matches 268; Conservative 65; Mismatches 161; Indels 16; Gaps 6;

QY 53 AGDEVQTLSSVRRAPTGP-PGTPGTCGAATAKPGCAQGP-----KPASLGRGRGA-A 104
DB 2 AGSDTAPFLSQADDDPDGPGFPGTGTGNGPKSELEVPDCEGLQRTIGLSRGSHTLI 61

QY 105 AAILSLGNVLYDRYTVAGVLLDIOQHFGVKDRGAGLLQSVFCSFMVAAPFVGLGR 164
DB 62 VVVLICYINLNYMDRFTVAGVLTIDIEQFFNIGDSSGLTQTFVFSYVWVAPVFGVGLGR 121

QY 165 FNRKVLSCGIFPWSAVTFSSFPQOYFWLLVSLRGLVIGEASYSTIAPTIGDLFTK 224
DB 122 YNRKYLKMGCGIAPFWSLVTLGSSFPICGEHFHLLLTGLVGVGEASYSTIAPTILDLFA 181

QY 225 NTRTLMLSVYFAIPGLSGVLYTGSVVKQAGDHWALRVSPVLGMITGTLILVLPAT 284
DB 182 DQSRMLSIIFYFAIPVGSGLGIYAGSKVMDAGDHWALRVTPGLGVVPLLVVREP 241

QY 285 KRGHADQ-LGDQLKARTSWLRDMKALIRNSVYVFFSLATSAYSFATGALGMWIPVYLRHA 343
DB 242 PRGAVEHSDLPPLNFTSNWADLALARNPFSVLSSLGFTSVAFVFGSLALWAPFLRS 301

QY 344 QVQKTAETC-NSPPCGAKDSLIFGAICTFTGFLGVVTVGAGATRCWLKTRQADPLVCV 402
DB 302 RVVLGETPPCLPGDCSSSDSLIFGLITCLTGLVGLGVGEISRLRHSNPRADPLVCV 361

QY 403 GMLGSAIFCLIFVAAKSIVGAYICIFVGETLLFSNWAITADILMYVVIPTTRATAVAL 462
DB 362 GLUGSAPFLFLSLACARGSVATYFIFIGETLLSNMVAIVADILLYVVIPTTRATAEAF 421
```

```

Db 362 GLLGSAFFLALACARGSIATYIFIFIGETLLSMNWAIVADILLYVVIPTRESTAEAF 421
Qy 463 QSFTSHLLGDAGSPYLIGFISDLIRQSTKYDPLWELFSLGVALMCLCPVVVLGGMFLAT 522
Db 422 QIVLSHLLGDAGSPYLIGLISDLRLRRSPWPSFLSEFRALQFSLMCLCAFVGALGAAFLGT 481
Qy 523 ALFFVSDRARAEOHLGERRAGVRVVRHQRGP 552
Db 482 AMFIEDRRRAQLHV-----QGLLHESGP 505

RESULT 8
Q9EQKO PRELIMINARY; PRT; 528 AA.
AC Q9EQKO;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Spinster-like protein.
GN Name=2210013K02R1k;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21238517; PubMed=11340170;
RA Nakano Y., Fujitani K., Kurihara J., Ragan J., Usui-Roki K.,
RA Shimoda L., Lukacsovich T., Suzuki K., Sezaki M., Sano Y., Ueda R.,
RA Awano W., Kaneda M., Umeda M., Yamamoto D.;
RT "Mutations in the novel membrane protein spinster interfere with
RT programmed cell death and cause neural degeneration in Drosophila
RT melanogaster.";
RL Mol. Cell. Biol. 21:3775-3788(2001).
RC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; AF212372; AAG33831.1; -;
DR MGD; MGI:1920908; 2210013K02R1k.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; F:transport; IEA.
DR InterPro; IPR000566; Lipocin_cytFABP.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PS00213; LIFOCALIN; UNKNOWN_1.
DR PROSITE; PS50850; MFS; 1.
KW Transmembrane.
SQ SEQUENCE 528 AA; 56860 MW; 85D4029B9B69827C CRC64;

Query Match 43.8%; Score 1278; DB 2; Length 528;
Best Local Similarity 52.2%; Pred. No. 3.2e-70;
Matches 266; Conservative 64; Mismatches 164; Indels 16; Gaps 6;

Qy 53 AGDEVQTLGSGVRRAPTPGP-FPTGTPGCAATAKPGGAQGP-----KPASLGRGRGA-A 104
Db 2 AGSDTAPFLSQADDDPGAPGHPGLPGPMGNPKSGELEVPDCEGLQRTGLSRGHSILI 61
Qy 105 AAILSLGNVLNLDRTYVAGVLLDITQHPGVKDRAGLGLQSFFICSFWVAAPIFGYLGDR 164
Db 62 VVVLICYINLLNMDRFTVAGVLTDEQFNIGDGTGLIQTIVFISYVNLAPVFGYLGDR 121
Qy 165 ENRKVLSCGIFFWASVTPSSSFIPOQYFWLLVLSRGLVIGGEASYSYTIPTIIGDLPTK 224
Db 122 YNRKVFMCGGTAFWSLVLTGSSSFIPIREHFWLLFLTRGMVGVEASYSYTIPTIADPFVA 181
Qy 225 NRTLMLSVFFPAIPGLSGGLGYITGSSVKQAAGDWHWALRSPVLGMITGTLLILVPAT 284
Db 182 DQSRMLSIFFYPAIPVSGSLGYIAGSKVKDVGADWHWALRVTGGLVLAELLFLVWQEP 241
Qy 285 KGHADQ-LGDLQKARTSWLRDMKALINRNSVYFSSLSATSAVSPATGALGMWIPLYLHRA 343
Db 242 PRGAVERHSGSPPLSPTSNWADLKALARNPSPVLSLGLFTSVAFVTGSLAWAPAFLLRS 301

```

```

Qy 344 QVYQKTAETC-NSPPPCGAKDSLIFGATICTGFLGVVVTGAGATRCWLKLTQRADPLVCAV 402
Db 302 RVVLGETPCLPGDSCSSDSLIFGLITCLTGVGLGMEISRRLLRFRNPADPLVCAA 361
Qy 403 GMLGSAIFICLIIVAAKSSIVGAYICIFVGETLLFNSWAITADILMYVVIPTRRATAVAL 462
Db 362 GLLGSAFFLALACARGSIATYIFIFIGETLLSMNWAIVADILLYVVIPTRESTAEAF 421
Qy 463 QSFTSHLLGDAGSPYLIGFISDLIRQSTKYDPLWELFSLGVALMCLCPVVVLGGMFLAT 522
Db 422 QIVLSHLLGDAGSPYLIGLISDLRLRRSPWPSFLSEFRALQFSLMCLCAFVGALGAAFLGT 481
Qy 523 ALFFVSDRARAEOHLGERRAGVRVVRHQRGP 552
Db 482 AMFIEDRRRAQLHV-----QGLLHESGP 505

RESULT 9
Q8UVB7 PRELIMINARY; PRT; 506 AA.
AC Q8UVB7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Not really started.
GN Name=spinl; Synonym=nr8;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21826209; PubMed=11836794;
RA Young R.M., Marty S., Nakano Y., Wang H., Yamamoto D., Lin S.,
RA Allende M.L.;
RT "Zebrafish yolk-specific not really started (nr8) gene is a vertebrate
RT homolog of the Drosophila spinster gene and is essential for
RT embryogenesis.";
RL Dev. Dyn. 223:298-305(2002).
RC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; AF465772; AAL69987.1; -;
DR ZFIN; ZDB-GENE-020228-1; spinl.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; F:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PS50850; MFS; 1.
KW Transmembrane.
SQ SEQUENCE 506 AA; 54962 MW; 033B0DE912CC1ED2 CRC64;

```

```

Query Match 43.4%; Score 1266.5; DB 2; Length 506;
Best Local Similarity 56.6%; Pred. No. 1.6e-63;
Matches 252; Conservative 66; Mismatches 124; Indels 3; Gaps 2;

Qy 107 ILSLGNVLNLDRTYVAGVLLDITQHPGVKDRAGLGLQSFFICSFWVAAPIFGYLGDR 166
Db 53 VLYCYINLLNMDRFTVAGVLPDIEHFGIGDGTGSLGLQTVFICSMFLAPLFGYLGDRYN 112
Qy 167 RKVLSCGIFFWASVTPSSSFIPOQYFWLLVLSRGLVIGGEASYSYTIPTIIGDLPTK 226
Db 113 RKLIMCVGIFFWSVTLASSFICKDHPWALLTRGLVGVGEASYSYTIPTIADLFVKEK 172
Qy 227 RRLMLSVFFPAIPGLSGGLGYITGSSVKQAAGDWHWALRSPVLGMITGTLLILVPATKR 286
Db 173 RTNMLSIFFYPAIPVSGSGMYIVGSKVDTVAKDWHWALRVTGGLLAVFLMLVYQBPKR 232
Qy 287 GHADQLGDLQKARTSWLRDMKALINRNSVYFSSLSATSAVSPATGALGMWIPLYLHRAQV 346
Db 233 GATEAHPHTLHRTSWLADMKALCRNPSPFILLSTFGTAVAFVTGSLAWAPAFLLRAGVF 292

```







QY 464 STSHLLGAGSPYLIGFTSDLIROSTKDSPIWEFLSLGVALMCLCFVVVLGMPFLATA 523  
 Db 371 IVLSHLLGAGSPYLIGFTSDLIROSTKDSPIWEFLSLGVALMCLCFVVVLGMPFLATA 430  
 QY 524 LFFVSDRAEAOHLGE--RRAGVR-----VVHORG 551  
 Db 431 IFIEADRRRAQLHVQGLLHAGSTDDRIIVVPORG 464

## RESULT 13

Q9D232 PRELIMINARY; PRT; 514 AA.  
 AC Q9D232;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
 DE Mus musculus adult male bone cDNA, RIKEN full-length enriched library,  
 DE clone:9830002117 product:weakly similar to SPINSTER-LIKE PROTEIN.  
 GN Name=9830002117Rik;  
 OS Mus musculus (Mouse);  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Bone;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning."  
 RL Meth. Enzymol. 303:19-44 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Bone;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690 (2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Bone;  
 RX The FANTOM Consortium;  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573 (2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Bone;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes."  
 RL Genome Res. 10:1617-1630 (2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Bone;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer."  
 RL Genome Res. 10:1757-1771 (2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Bone;  
 RX Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,

RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai T.,  
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Muramatsu M., Hayashizaki Y.;  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 DR EMBL; AK020683; BAB32173.1; -;  
 DR MGD; MGI:1924827; 9830002117Rik.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR000566; Lipoclin\_cytFABP.  
 DR InterPro; IPR007114; MFS.  
 DR InterPro; IPR00215; Prot\_inh\_serpin.  
 DR InterPro; IPR005828; Sub transporter.  
 DR Pfam; PF00083; Sugar tr; 1.  
 DR PROSITE; PS00213; LIPOCALIN; UNKNOWN\_1.  
 DR PROSITE; PS00850; MFS; 1.  
 DR PROSITE; PS00284; SERPIN; UNKNOWN\_1.  
 KW Transmembrane.  
 SQ SEQUENCE 514 AA; 55456 MW; DCCBACFEC340CE74 CRC64;

Query Match 39.0%; Score 1139; DB 2; Length 514;  
 Best Local Similarity 43.5%; Pred. No. 1e-61;  
 Matches 243; Conservative 80; Mismatches 163; Indels 72; Gaps 7;  
 QY 2 MCLECAAAAGGAEEEDAEERRRRRGAQRGAGSGCGGARGAGGAGVSAAGDEVQTLS 61  
 Db 1 MSTELCKPQTGCPQSLS-----QGQYGLASGTC----- 32  
 QY 62 GSVRRAPTPGPTGPTGCAATAGKGAQPKPASLGRG-AAAAIISLGNVLNLDY 120  
 Db 33 -----LPPST-----PVPNSLPRWRAYLAAAVLCYINLLNMMWF 67  
 QY 121 TVAGVLLDIQQHFGVKRGAGLQSVFICSFVVAPIFGYLGDRNRKVVILSCGIFPWSA 180  
 Db 68 IIPGVLLDVQYFHISDSHAGLQTVFISCLLSAPVFGYLGDRYNRKAILISFGILLWSG 127  
 QY 181 VTFSSSFIPQOYFWLLVLSRGLVGIGEASYSYTIPTIIGDLFTKNTRLMLSVFYFAIPL 240  
 Db 128 AGLSSSFISYQYSWLFFLSRGFVGTAASYSTIAPTIVLGDLFVKQRTCALAVFIPIV 187  
 QY 241 CSGLYITGSSVKQAAGDMHWALRVSPVLGMITGTLLILVLPATKRGHADQLGD-OLKA- 298  
 Db 188 GSGLYVLGTVTAELTGNRWALRMLPCLDAMALALLLVDPVPRGAERKQGEVAVRAP 247  
 QY 299 RTSWLRDMKALIRNSYVFSSLSATSAVSFATGALGMWIPLYLHRAQVVKTAETCNSPCC 358  
 Db 248 RSSMCEVDVYLRNWSFVSTLGTATIAFTVGTGAFVFNAPKFLFEARVHGLQPCFQBC 307  
 QY 359 GAKDSLIFGALTCTGFLGVVTGAGATRCWKRTQADPLVCAGMGLSALFICILFVAA 418  
 Db 308 HSQDSLIFGALTATVATIGVLMGLAESRYKKNVPRAEPLICASSLPATACLYLALILA 367  
 QY 419 KSSIVAGYICIFVGETLLFSNWAITADILMYVVIPTTRATAVALQSFTHLLGDAGSPYL 478  
 Db 368 SRTLLASVFLALGELLISCNWAVADILLSVWVPRCRGTAEALQITVAHVILGDAGSPYL 427  
 QY 479 IGFTSDLIROSTKDSPIWEFLSLGVALMCLCFVVVLGMPFLATLFPVSDRAEAOHLG 538  
 Db 428 TGLISSVLQAEERPDPYSLQHFSLQHSFLCCAFIVLGGGFFLLTALHLEKQARQ--- 484  
 QY 539 ERRAGVRVHQRGPQGT 556  
 Db 485 -----PQKGT 489

RESULT 14

```
Q71RB5
ID Q71RB5 PRELIMINARY; PRT; 420 AA.
AC AAQ15259;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE PP2030.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang Y.; Zhou X.M.; Zhang P.P.; Jiang H.Q.; Qin W.X.; Zhao X.T.;
RA Wan D.F.; Gu J.R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF370423; AAQ15259.1; -.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS0850; MFS; 1.
SQ SEQUENCE 420 AA; 45301 MW; 5771BC5C252CD048 CRC64;
Query Match 38.2%; Score 1115; DB 2; Length 420;
Best Local Similarity 56.6%; Pred. No. 2.4e-60;
Matches 231; Conservative 50; Mismatches 119; Indels 8; Gaps 4;
152 MVAAPIFGVLGDRFNKRVILSCGIFFWSAVTFSSSFIPOQYFWLIVLSRGLVGEASYS 211
1 MVLAPVGVGLGDRYNRKYLKMGCGIAFWLSVLTGSSSFIPIGEHFWLLLTTRGLVGVGEASYS 60
212 TIAPTIIGDLFTKNTRTMLSVFYFAIPGLGSLGYITGSSVKQAAGDMHWALRVSPVLGM 271
61 TIAPTLIADLFVADQSRMLSIIFYFAIPVSGSLGYIAGSKVDMAGDMHWALRVTPGLGV 120
272 ITGTILILVPATKRGHADQDQDKAR-TSWLRDMKALIRNRSYVFSSLSATSAVSPTG 330
121 VAVLLFLVVRPPRGAVRHSDLPPLNPTSMWADLRALARNPSFVLSLGLFTAVAPVTG 180
272 ITGTILILVPATKRGHADQDQDKAR-TSWLRDMKALIRNRSYVFSSLSATSAVSPTG 330
121 VAVLLFLVVRPPRGAVRHSDLPPLNPTSMWADLRALARNPSFVLSLGLFTAVAPVTG 180
331 ALGWMIPLYLHRAQVVKTAETC-NSPPCGAKDSLIFGAICTCTGFLGVVGTGATRWCR 389
181 SLALWAPAFLLRSRVVLGETPPCLPGDSCSSDSLIFGLITCLTGLVGLVGEISRRLR 240
390 LKTQADPLVCAGVGLGSAIFICLIFVAAKSSIVGAYICIFVGETLLFSNWAITADILMY 449
241 HSNPRADPLVCATGLLGSAPFLFLSLACARGSIVATYIFIFIGETLLSMNWAIVADILY 300
450 VVIPTRRATAVALQSFTSHLLGAGSPYLIIGFISDLIRQSTKDSPLWEFLSLGYALMLCP 509
301 VVIPTRRATAVALQSFTSHLLGAGSPYLIIGFISDLIRQSTKDSPLWEFLSLGYALMLCP 509
510 FVVVLGGMFFLATLFFVSDRARAEOHLGE--RRAGVR----VVHQRG 551
361 FVGALGGAFLGTATFIADRRRAQLHVQGLLHEAGSTDDRIVVPPQRG 408
```

## RESULT 15

```
AAQ15259
ID AAQ15259 PRELIMINARY; PRT; 420 AA.
AC AAQ15259;
DT 02-MAR-2004 (Tremblrel. 27, Created)
DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)
DT 02-MAR-2004 (Tremblrel. 27, Last annotation update)
DE PP2030.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang Y.; Zhou X.M.; Zhang P.P.; Jiang H.Q.; Qin W.X.; Zhao X.T.;
RA Wan D.F.; Gu J.R.;
RL "Novel human cDNA clones with function of inhibiting cancer cell growth."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; AF370423; AAQ15259.1; -.
SQ SEQUENCE 420 AA; 45301 MW; 5771BC5C252CD048 CRC64;
Query Match 38.2%; Score 1115; DB 2; Length 420;
Best Local Similarity 56.6%; Pred. No. 2.4e-60;
Matches 231; Conservative 50; Mismatches 119; Indels 8; Gaps 4;
152 MVAAPIFGVLGDRFNKRVILSCGIFFWSAVTFSSSFIPOQYFWLIVLSRGLVGEASYS 211
1 MVLAPVGVGLGDRYNRKYLKMGCGIAFWLSVLTGSSSFIPIGEHFWLLLTTRGLVGVGEASYS 60
212 TIAPTIIGDLFTKNTRTMLSVFYFAIPGLGSLGYITGSSVKQAAGDMHWALRVSPVLGM 271
61 TIAPTLIADLFVADQSRMLSIIFYFAIPVSGSLGYIAGSKVDMAGDMHWALRVTPGLGV 120
272 ITGTILILVPATKRGHADQDQDKAR-TSWLRDMKALIRNRSYVFSSLSATSAVSPTG 330
121 VAVLLFLVVRPPRGAVRHSDLPPLNPTSMWADLRALARNPSFVLSLGLFTAVAPVTG 180
331 ALGWMIPLYLHRAQVVKTAETC-NSPPCGAKDSLIFGAICTCTGFLGVVGTGATRWCR 389
181 SLALWAPAFLLRSRVVLGETPPCLPGDSCSSDSLIFGLITCLTGLVGLVGEISRRLR 240
390 LKTQADPLVCAGVGLGSAIFICLIFVAAKSSIVGAYICIFVGETLLFSNWAITADILMY 449
241 HSNPRADPLVCATGLLGSAPFLFLSLACARGSIVATYIFIFIGETLLSMNWAIVADILY 300
450 VVIPTRRATAVALQSFTSHLLGAGSPYLIIGFISDLIRQSTKDSPLWEFLSLGYALMLCP 509
301 VVIPTRRATAVALQSFTSHLLGAGSPYLIIGFISDLIRQSTKDSPLWEFLSLGYALMLCP 509
510 FVVVLGGMFFLATLFFVSDRARAEOHLGE--RRAGVR----VVHQRG 551
361 FVGALGGAFLGTATFIADRRRAQLHVQGLLHEAGSTDDRIVVPPQRG 408
Search completed: October 23, 2004, 13:14:07
Job time : 115 secs
```